

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

# OM protein - nucleic search, using frame\_p2n model

Run on: December 24, 2004, 20:52:22 ; Search time 3665 Seconds  
(without alignments)  
2498.455 Million cell updates/sec

Title: US-09-743-885A-1  
Perfect score: 1353  
Sequence: 1 MAGLADDSFIFGLGNIV.....NSGNALQSMKDPRLRTSK 265

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 32822875 seqs, 18219865908 residues  
Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0  
Maximum DB seq length: 200000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODE=frame+g2n.model -DRV=xlp  
-Q/cgn2.1/USPTO.spool.p/US09743865/runat\_23122004.165259.9638/app\_query.fasta\_1.455  
-DB=EST -OPMT=fasta -SUFFIX=rest -MINMATCH=0.1 -LOOPEXT=0 -LIST=45  
-UNITS-bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=pco -NOR=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09743865\_@CEN\_1\_1\_6425\_@runat\_23122004.165259.9638 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:.\*  
1: gb\_est1:.\*  
2: gb\_est2:.\*  
3: gb\_hc:.\*  
4: gb\_est3:.\*  
5: gb\_est4:.\*  
6: gb\_est5:.\*  
7: gb\_est6:.\*  
8: gb\_gsa1:.\*  
9: gb\_gsa2:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	796.5	58.9	707	1	AJ558666
2	713.5	52.7	935	6	CB981330
3	706.5	52.2	747	6	CB970796
4	694	51.3	843	7	CB970796
5	691.5	51.1	771	6	CB979874
6	680.5	50.3	927	3	CNS0A9R7
7	677.5	50.1	765	7	CF445200
8	668.5	49.4	731	1	AJ796299
9	660.5	48.8	665	5	BU011719

10	656	48.5	872	7	CO125949	CO125949 GR_EB09L
11	656	48.5	881	7	CO123087	CO123087 GR_EB05A
12	655.5	48.4	865	7	CO125993	CO125993 GR_EB09M
13	655.5	48.4	888	7	CO121875	CO121875 GR_EB03E
14	654.5	48.4	865	7	CO121803	CO121803 GR_EB03C
15	653.5	48.3	858	7	CO090096	CO090096 GR_EA090
16	652.5	48.2	796	7	CO102085	CO102085 GR_EB002
17	652.5	48.2	796	7	CO110654	CO110654 GR_EB004
18	652.5	48.2	832	7	CO108163	CO108163 GR_EB003
19	651	48.1	737	7	CO112381	CO112381 GR_EB004
20	651	48.1	778	7	CO120905	CO120905 GR_EB024
21	650.5	48.1	1446	3	AY105872	AY105872 Zea mays
22	650	48.0	788	7	CF443959	CF443959 EST680304
23	646.5	47.8	715	7	CN125871	CN125871 RH0H1.13
24	646.5	47.8	750	4	BG840650	BG840650 MEST14-F0
25	644.5	47.6	745	4	BI422022	BI422022 EST532688
26	644	47.6	881	7	CO101149	CO101149 GR_EB002
27	640.5	47.3	737	4	BM113166	BM113166 EST560702
28	639.5	47.3	684	5	BO994945	BO994945 QGFBJ23.Y
29	635.5	47.0	786	7	CO117282	CO117282 GR_EB019
30	634	46.9	727	4	BI421916	BI421916 EST532582
31	632.5	46.7	779	7	CO099006	CO099006 GR_EA23K
32	632.5	46.7	842	7	CO087535	CO087535 GR_EA06A
33	632.5	46.7	872	7	CO094841	CO094841 GR_EA17C
34	631.5	46.7	753	7	CO099630	CO099630 GR_EA24K
35	629.5	46.5	627	6	CA137196	CA137196 SCCCR7200
36	629.5	46.5	793	5	BO118756	BO118756 EST604319
37	629	46.5	885	7	CK288429	CK288429 EST751151
38	629	46.5	942	7	CK294790	CK294790 EST751504
39	629	46.5	968	7	CK286610	CK286610 EST749332
40	628.5	46.5	847	7	CO130491	CO130491 GR_EB33E
41	627.5	46.4	1127	7	CK161576	CK161576 PGAS01414
42	623.5	46.1	774	7	CO131542	CO131542 GR_EB430
43	623.5	46.1	1101	3	CNS0A181	EX830132 Arabidops
44	623	46.0	739	2	AW441207	AW441207 EST310603
45	623	46.0	757	4	BG592948	BG592948 EST491626

## ALIGNMENTS

RESULT 1  
AJ558666  
LOCUS  
DEFINITION  
AJ558666 Antirrhinum majus whole plant Antirrhinum majus cDNA clone  
018.1 08 c19, mRNA sequence.

ACCESSION  
AJ558666  
VERSION  
AJ558666.1 GI:31661238  
KEYWORDS  
EST.  
SOURCE  
Antirrhinum majus (snapdragon)  
ORGANISM  
Antirrhinum majus

REFERENCE  
1 (bases 1 to 707)  
Zachgo, S., Stueber, K., Saedler, H., Sommer, H. and Schwarz-Sommer, Z.  
Antirrhinum EST collection  
Unpublished (2003)  
Contact: Schwarz-Sommer Z

COMMENT  
Molekulare Pflanzen-genetik  
MPI fuer Zuechtungs-forschung  
Carl-von-Linne Weg 10, D-50829, Germany.  
FEATURES  
Source  
1. 707  
/organism="Antirrhinum majus"  
/mol\_type="mRNA"  
/db\_xref="taxon:4151"  
/clone="018.1 08 c19"  
/tissue\_type="whole plant"  
/clone\_id="Antirrhinum majus whole plant"

## ORIGIN

Alignment Scores:

Pred. No.: 2.56e-82 Length: 707  
 Score: 796.50 Matches: 152  
 Percent Similarity: 83.49% Conservative: 30  
 Best Local Similarity: 69.72% Mismatches: 33  
 Query Match: 58.87% Indels: 3  
 Gaps: 2

US-09-743-885a-1 (1-265) x AJ558666 (1-707)

QY 4 LeuArg1aAspAspLeuSerPhe1lePheGlyLeuGluGlyAen1leValSerPheMet 23  
 Db 11 TTAATGCTCAAGCTTGGCTTTTATATTTGGCTATTTGGTAACATATATCTTTCTG 70  
 QY 24 ValPheLeuAlaProValProThrPheTyrLys1leTyrLysArgLysSerSerGluGly 43  
 Db 71 GTATTTTGGACACGAGCCAACTTTTACGATTTATAGACGAATATCATGAAAGG 130  
 QY 44 TycGlnAla1lePheTyrMetValAlaLeuPheSer1leGlyLeuLeuTyrTyrAla 63  
 Db 131 TTTGAGCGATACCGTATTCGTTGCTTTAGTCTCTTACTGCTATACTATGCA 190  
 QY 64 TyrLeuArgLysAsnAlaTyrLeu1leValSer1leAsnGlyPheGlyCybAla1leGlu 83  
 Db 191 GTTATTAGAGCATGCTTATATATGTTATGATTCATAGATTTGATGATTCATTGAA 250  
 QY 84 LeuThrTyrLysSerLeuPheLeuPheTyrAlaProArgLysSerLys1lePheThrGly 103  
 Db 251 GCAGTTTACCATCAATCTACTTATATATCATCCAAAGTCTTAAGTTTCACGCGTG 310  
 QY 104 TripleMetLeuLeu---GluLeuGlyAlaLeuGlyMetValMetPro1leThrTyrLeu 122  
 Db 311 AGATTGATGACTATTCATGATGAGGTTTATGAGTGTGATGCGAGTCTTTGCTA 370  
 QY 123 LeuAlaGluGlySerHisArgValMet1leValGlyTyr1leCybAla1leAsnVal 142  
 Db 371 GCTGTCACGCTGCAAAAAGGCTCTCTGCTGGAGATGAGATGCTTCATTTACGCTT 430  
 QY 143 AlaValPheAla1leProLeuSer1leMetArgGlnVal1leLysThrLysSerValGlu 162  
 Db 431 GCCGTTTTCGCTCTCTTATGACATATGAGCGAGTCATAGAACCAAAAAGCGTAGAG 490  
 QY 163 PheMetProPheThrLeuSerLeuPheLeuThrLeuCybAlaThrMetTyrPhePheTyr 182  
 Db 491 TTCATGCGATTCACGCTATCTTCTCTCCACCTTTGCTACAAATGATGTTCTTAT 550  
 QY 183 GlyPhePheLysLysAspPheTyr1leAlaPheProAsn1leLeuGlyPheLeuPheGly 202  
 Db 551 GGGTCTCTTTGTTAGGACCCCTACATCGCTTGGCAATGCTTGGGATTTCTTATTTGGG 610  
 QY 203 1leValGlnMetLeuLeuTyrPheValTyrLysAspSerLysArg1leAspAsp 220  
 Db 611 ATTGCTCAAAATGATTGTTATTCATCTACAAAGAACCCCAA-----GACGAT 658

RESULT 2  
 CB981330 935 bp mRNA linear EST 01-MAY-2003  
 LOCUS DEFINITION CAB70004.1IaF\_G09 Cabernet Sauvignon Berry Post-Veraison - CAB7  
 ACCESSION CB981330  
 VERSION CB981330.1 GI:30304536  
 KEYWORDS EST.  
 SOURCE Vitis vinifera  
 ORGANISM Vitis vinifera  
 Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; Vitaceae; Vitis.  
 1 (bases 1 to 935)  
 Goes de Silva,F., Iandolo,A., Lim,H., Baek,J., Leslie,A., Xu,J.,  
 Jones,K. and Cook,D.  
 Expressed sequence tags from cabernet sauvignon berries at various  
 developmental stages  
 JOURNAL Unpublished (2003)  
 COMMENT Contact: Douglas Cook, PhD

CAB7 Genome Facility  
 UC Davis, Plant Pathology  
 One Shields Ave, Davis, CA 95616, USA  
 Tel: 530 754 6561  
 Fax: 530 754 6617  
 Email: drcocook@ucdavis.edu  
 Seq primer: ACGGTACCGCATATGCC.  
 Location/Qualifiers

#### FEATURES

##### SOURCE

1. 935

/organism="Vitis vinifera"

/mol\_type="mRNA"

/cultivar="Cabernet Sauvignon"

/db\_xref="taxon:29760"

/clone="CAB70004.1IaF\_G09"

/sex="Hermaphrodite"

/dev\_stage="Post-Veraison, 18-19 drix"

/lab\_host="DH5alpha"

/clone\_lib="Cabernet Sauvignon Berry Post-Veraison - CAB7"

/note="Organ: Berry; Vector: pDNR; Site: 1: SfiI; Site 2: SfiI; CAB7 is a cDNA library of Cabernet Sauvignon clone berries. Samples were collected post-veraison from field-grown vines during stage III of berry growth at 89 days after full bloom. Berries soluble solid concentration ranged between 18-19 drix. Sampled vines were located at the University of California, Davis, Experimental vineyard. cDNAs were made by oligo-dT priming and directl onally cloned. 5' and 3' adaptors were used in cloning as follows: 5'-AAGCAGTGGTCAAGAGAGTGGCCATTACGCGCGG-3' and 5'-ATTCTAGAGCGGAGGCGCGAGATG-dT(30)NN-3'. Library was constructed using the Clontech Creator SMART kit and size-selected to contain the 0.5-3 kb size fraction."

#### ORIGIN

##### Alignment Scores:

Pred. No.: 1.85e-72 Length: 935  
 Score: 713.50 Matches: 138  
 Percent Similarity: 69.53% Conservative: 56  
 Best Local Similarity: 49.46% Mismatches: 46  
 Query Match: 52.73% Indels: 39  
 Gaps: 4

US-09-743-885a-1 (1-265) x CB981330 (1-935)

QY 1 MetAlaGlnLeuArg1aAspAspLeuSer-----Phe 11  
 Db 43 CTGGCCACCTCCGACGTAAGATGCGCTGTTTCCCATTCACCATCTCTGGTTT 102  
 QY 12 1lePheGlyLeuGluGlyAen1leValSerPheMetValPheLeuAlaProValProThr 31  
 Db 103 ATCTTGCGCATCTTAGGTAACTCATCTCATATGATGATGATCTTCTTACGCTGACT 162  
 QY 32 PheTyrLys1leTyrLysArgLysSerSerGluGlyTyrGlnAla1leProTyrMetVal 51  
 Db 163 TTCTACCAATTTTACAAAGAAATCACATCGAAGGTTTATGATGATGATGATGATGAT 222  
 QY 52 AlaLeuPheSer1leGlyLeuLeuTyrTyrAlaTyrLeuArgLysAsnAlaTyrLeu 71  
 Db 223 GCATTTATTCAGTCTATGCTTGGATATATCTTCTTAAACATGATGATGATGATGAT 282  
 QY 72 1leValSer1leAsnGlyPheGlyCybAla1leGluLeuThrTyrLysSerLeuPheLeu 91  
 Db 283 CTCATCCATTCACCAATCCGTTGGCTGCTGATGAGCCGATCAATGTTATGTTCTT 342  
 QY 92 PheTyrAlaProArgLysSerLys1lePheThrGlyTyrLeuMet---LeuLeuGluLeu 110  
 Db 343 GTGTATGACCGAAGAGGCTTAGATCACGACTGTGAATCTGTTTCTTATGATATATC 402  
 QY 111 GlyAlaLeuGlyMetValMetPro1leThrTyrLeuLeuAlaGluGlySerHisArgVal 130  
 Db 403 TGGGATGCGTTCCATCGCTGCTTAACCTCTCTAGCTGAAGCGGTACCGGCTC 462  
 QY 131 Met1leValGlyTyr1leCybAla1leAsnValAlaValPheAla1leProLeuSer 150

Db	463	CGTATTTCTGGATGGGTTTGTCTGCTGTTCTTCGCTAAACGGTTTCTTCGACACCCCTTTCG	522
Qy	151	lIleMetArgInValIleIleYThrIysSerValGluPheMetProPheThrIleuSerIeu	170
Db	523	ATCATGAGACAAAGTATCTCGAACCAAGACGCTGTGAGTCAATGCCATTTCTTCTTACCTTT	582
Qy	171	PheIeuThrIeuCysAlaThrMetTyrPhePheTyrGlyPhePheIleValGlnMetIleuTyrPhe	190
Db	583	TTCTTCACATTGAGTGTCTGTAAATGTGGTCTTTTATGGTCTGATGCTCAAGACCTTTAC	642
Qy	191	lIleAlaPheProAsnIleIleuGlyPheIeuPheGlyIleValGlnMetIleuTyrPhe	210
Db	643	ATTGCTGGTCTCAAAACATCTGGGCTGTCTCTTCGGGATTTGTTACAGATGTCCTCTACTTA	702
Qy	211	ValTyrIleValAspSerIysArgIle---AspAspGlyIysSerAspProValArgIle---	228
Db	703	ATCTACAGAAACAGAGAAAGATCTCGAAGAACGAAATTGCTGAACTATCCGACAA	762
Qy	228	-----	228
Db	763	ATCATGTACGCTTGTGAAGCTTAGCAACATGGTGTGCTCAGAAAGTAACTAACAAG	822
Qy	229	-----AlaThrIysSerIysGlyIleValGluIleIleIle	240
Db	823	CAGCAGACGATGAAAGACATGCGCACACAGAGAAACAAGGCTTGAGTCAATAGTT	879
RESULT 3			
CB970796			
LOCUS	CB970796	747 bp	mRNA linear EST 30-APR-2003
DEFINITION	CAB10004_Ira_Fa_F04 Cabernet Sauvignon flower pre-bloom - CAB1		
ACCESSION	Vitis vinifera cDNA clone CAB10004_Ira_Fa_F04 5', mRNA sequence.		
VERSION	CB970796		
KEYWORDS	CB970796.1 GI:30253245		
SOURCE	EST.		
ORGANISM	Vitis vinifera		
	Vitis vinifera		
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
AUTHORS	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;		
	rosoids; Vitaceae; Vitis.		
	1 (bases 1 to 747)		
	Goes da Silva,F., Jandolino,A., Lim,H., Baek,J., Leslie,A., Xu,J.,		
	Jones,K. and Cook,D.		
TITLE	Expressed sequence tags from Vitis vinifera 'Cabernet sauvignon'		
	berries at various developmental stages		
	Unpublished (2003)		
JOURNAL	Contact: Douglas Cook, PhD		
COMMENT	CARS Genome Facility		
	UC Davis, Plant Pathology		
	One Shields Ave, Davis, CA 95616, USA		
	Tel: 530 754 6561		
	Fax: 530 754 6617		
	Email: drcook@ucdavis.edu		
FEATURES	Seq primer: ACGTACCGGACATATGCC.		
SOURCE	Location/Qualifiers		
	1..747		

cloning as follows:  
5'-AAGCAGTGTATCAGCCAGAGTGGCCATTACGGCCGGG-3' and  
5'-ATTCTAGAGCCGAGGGGGCCGACATG-dt(30)NN-3'. Library was  
constructed using the Clontech Creator SMART kit and  
size-selected to contain the 0.5-3 kb size fraction."

cloning as follows:  
 5'-AAGCAGTGTGTTACAGCCAGATGGCCATTACGCCCGG-3' and  
 5'-ATTCTAGAGCGCCGACGGCGCGGACATG-dt(30)NN-3'. Library was  
 constructed using the Clontech Creator SMART kit and  
 size-selected to contain the 0.5-3 kb size fraction."

Alignment Scores:	Pred. No.:	8,95e-72	Length:	747
Score:	706.50	Matches:	133	
Percent Similarity:	78.54%	Conservative:	50	
Best Local Similarity:	57.08%	Mismatches:	39	
Query Match:	52.22%	Indels:	11	
DB:	6	Gaps:	3	

  

US-09-743-885A-1 (1-265) x CB970796 (1-747)	Query	1	MetaIaqlnleuArlgaIaaspAhpIeuser-----Phe	11
	Db	35	CTTCCGCCACCTCCCGAGTGAAGAGATGCGCTTGTTCACATTCACATCCCTTGGTATT	94
	Qy	12	IllephgIleuLeuEngIyaenIlleValSerPheMetValPheleuAlaProvalProthr	31
	Db	95	ATCTTTGGCATCTCAGGTAACCTATCTATCTATCACTGATGTAACCTTGCCTCCACCTG	154
	Qy	32	PheTrpLysIleTrpLysArgLysSerSerGluGlyTrpGlnAlaIleProTrpMetVal	51
	Db	155	TTTACCAAAATTTTCAAAAGGAATCCACGTAAGGGTTTCAGTCTGTTCCTTACGATG	214
	Qy	52	AlaIleuPheSerAlaGlyLeuLeuLeuTrpTrpAlaTrpLysArgLysAspAlaTrpLeu	71
	Db	215	GCATTATTCAGTGTGTAAGCTTTGGATATACATATGCTTCCTTAACACTGATGTA	274
	Qy	72	IlleValSerIleValEngIlePheGlyCysAlaIleGluLeuTrpTrpIleSerLeuPheLeu	91
	Db	275	CTCATCAACATCAACATCGGTGGCTGGGTGATTAAGACACACTACATGTTATGTTCTT	334
	Qy	92	PheTrpAlaProAlaGlySerLysIlePheTrpGlyTrpLeuMet---LeuLeuGluLeu	110
	Db	335	GTTGATTCACCGAAGAGGCTAGGATCAACGCTGTGAACTCGTTTTCTATATGATATC	394
	Qy	111	GlyAlaLeuEngIleMetValMetProIleThrTrpLeuLeuAlaGluGlySerHisArgVal	130
	Db	395	TGGCGATTCGGTTCATCTCTGCTCTTACCTCTCTCTTACCTGAAGGCGCTAACCGTCTC	454
	Qy	131	MetIleValGlyTrpIleCysAlaAlaIleAsnValAlaValPheAlaAlaProLeuSer	150
	Db	455	CGATTCTTGGATGGGTTGTGCTGCTGTTCCGTAAGCGTTTTCTTGCACCCCTTGGC	514
	Qy	151	IleMetArgGlnValIleLysThyLysSerValGluPheMetProPheThrLeuSerLeu	170
	Db	515	ATCATGAGACAAAGTTATCCGAACCAAGGCGTTAGTACACCAATTTCTTATCTTTT	574
	Qy	171	PheLeuThrIleuCysAlaIleMetTrpPheTrpGlyPhePheValYAspPheTrp	190
	Db	575	TTCTTCACATTTGAAGTGGCTGTAATGTTGTTCTTTATGGTCTGATGCTCAAGACATT	634
	Qy	191	IleAlaIlePheProAlaIleLeuGlyPheLeuPheGlyIleValGlnMetLeuLeuTrpHe	210
	Db	635	ATTCTGCTGCMAACATCCCTGGGCTGTGCTTGGGATTTGTTTCAGATGCTCTTACTTA	694
	Qy	211	ValTrpLysAspSerLysArgIle---AspArgGluLys	222
	Db	695	ATCTACAAAACACAGAAAGAAAGTTCTGGAACACAGAAA	733

  

RESULT 4	CF517565	843 bp	mRNA	linear	EST 09-SEP-2003
LOCUS	CF517565				
DEFINITION	CAP0004_IVF_A05 Virus vinfifera cv. cabernet sauvignon (Clone 8)				
	Petiole - C&P Vitis vinifera cDNA clone CAP0004_IVF_A05 5', mRNA				
	Sequence.				
ACCESSION	CF517565				
VERSION	CF517565.1				GI:34549333





/clone\_1lb="Cabernet Sauvignon Berry Post-Veraison - CAB7"  
 /note="Organ: Berry; Vector: pDNR; Site 1: SfiI; Site 2: SfiI; CAB7 is a cDNA library of Cabernet Sauvignon Clone 8 berries. Samples were collected post-veraison from field-grown vines during stage III of berry growth at 89 days after full bloom. Berries soluble solid concentration ranged between 18-19 brix. Sampled vines were located at the University of California, Davis, Experimental vineyard. cDNAs were made by oligo-dT priming and directly cloned. 5' and 3' adaptors were used in cloning as follows: 5'-AAGCAGTGTATCAAGCAGATGCGCATACGGCCGG-3' and 5'-ATCTAGAGCGCGAGCGCGCATG-dT(30)NM-3'. Library was constructed using the Clontech Creator SMART kit and size-selected to contain the 0.5-3 kb size fraction."

## ORIGIN

## Alignment Scores:

Pred. No.:	Score:	Percent Similarity:	Best Local Similarity:	Query Match:	DB:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
5.28e-70	691.50	77.68%	55.79%	51.11%	6	771	130	51	41	11	3

US-09-743-885a-1 (1-265) x CB979874 (1-771)

```

QY 1 MetAlaGlnLeuArgAlaAspLeuSer-----Phe 11
DB 34 CTGCGCCACCTCCACGCTGAAGAGATGCTTGTCCATTCACCTCTTGTTT 93
QY 12 IllePheGlyLeuLeuGlyAlaValSerPheMetValPheLeuAlaProValProThr 31
DB 94 ATCTTCGGCATCTCAGTAACTCCTCATCTCATGTGATGCTGCTCCACGCCACT 153
QY 32 PheTyrIleValIleTyrIleValSerSerGluGlyTyrGlnAlaIleProTyrMetVal 51
DB 154 TTCTACCAAAATTACAAAGAAATTCACATGAAGGGTTCAGTCTGTTCCCTCCTG 213
QY 52 AlaLeuPheSerIleGlyLeuLeuLeuTyrTyrAlaTyrLeuArgIleValAlaTyrLeu 71
DB 214 GCATTATTCAGTCTATGCTTTGGATATGCTTCTTCTTAACTGATGCTAAGCTT 273
QY 72 IleValSerIleLeuGlyPheGlyCysAlaIleGluLeuThrTyrIleSerLeuPheLeu 91
DB 274 CTATATCACCTCACTACCTCGTGGCTGCTATGACACGACATGTTTATGTTCTT 333
QY 92 PheTyrAlaProValSerIleValPheThrGlyTyrPheMet---LeuLeuGluLeu 110
DB 334 GTGTATGCACGGAAGAGGCTAGATCAGCATGTAACCTCGTTTCTTAAATGAAATATC 393
QY 111 GlyAlaLeuGlyMetValMetProIleThrTyrLeuLeuAlaGluGlySerHisArgVal 130
DB 394 TGGCGATTCGGTTCACCTGCTTAACTCTCTTCTTAACTGAAAGCGCTAACCGTGC 453
QY 131 MetIleValGlyTyrIleCysAlaAlaIleAsnValAlaValPheAlaIleProLeuSer 150
DB 454 CGATTCTTGGATGGGTTTGTCTGCTGCTTCTGCTTAAAGCTTTTCTTGACACCCCTTGC 513
QY 151 IleMetArgIleValIleValSerValGluPheMetProPheThrLeuSerLeu 170
DB 514 ATCATGAGCAAGATTATCCGACCAAGCGTTGATGATGCTGATGCTCAAGACTTTTAC 573
QY 171 PheLeuThrLeuCysAlaThrMetTyrPhePheTyrGlyPhePheIleValAspPheTyr 190
DB 574 TTCTTCATCATGATGCTGATGATGCTTTTATGATGCTGATGCTCAAGACTTTTAC 633
QY 191 IleAlaPheProAlaIleLeuGlyPheLeuPheGlyIleValGluMetLeuTyrPhe 210
DB 634 ATTCGCTGCTCAAAATCCATCCATCTTCTGCTGCGGATGTTTCTGATGCTCTACTTA 693
QY 211 ValTyrIleAspSerIleValSer---AspAspGluVal 222
DB 694 ATCTACGAAACAGAGAAAGTTCTGAAACAGAGAA 732

```

## RESULT 6

CNS0A9RJ 927 bp mRNA linear HTC 06-FEB-2004  
 LOCUS Arabidopsis thaliana Full-length cDNA Complete sequence from clone  
 DEFINITION G57FB532F04 of Flowers and buds of strain col-0 of Arabidopsis  
 thaliana (thale cress).

## ACCESSION

VERSION BX819171.1 GI:42468928  
 KEYWORDS HTC; GSUT cDNA  
 SOURCE Arabidopsis thaliana (thale cress)  
 ORGANISM Arabidopsis thaliana

## REFERENCE

AUTHORS Castelli,V., Aury,J.M., Jallou,O., Wincker,P., Clepet,C., Menard,M., Cruaud,C., Queller,C., Scarpelli,C., Schachter,V., Temple,G., Caboche,M., Weissenbach,J. and Salanoubat,M.  
 TITLE Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences: A Combined Approach to Evaluate and Improve Arabidopsis Genome Annotation

REFERENCE Unpublished  
 AUTHORS 2 (bases 1 to 927)  
 TITLE Genoscope.

JOURNAL Direct Submission  
 TITLE Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage ; BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr

## COMMENT

The sequences are based on single pass reads.  
 Life Technologies (a division of Invitrogen) members carried out full-length libraries construction ; Temple G.  
 Genoscope members carried out sequencing and annotation ; Castelli V., Aury J.M., Jallou O., Wincker P., Menard M., Cruaud C., Schachter V., Weissenbach J., Salanoubat M.  
 URV INRA : Clepet C., Caboche M.  
 Annotation is based on the June 2003 version of the Arabidopsis genome released by MIPS (Munich Information center for Protein Sequences). 5 prime and 3 prime are assembled with Phrap.  
 http://www.genoscope.cns.fr/externe/sequences/Banque\_Projet\_EF/Full length  
 http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.

## FEATURES

source location/Qualifiers  
 1..927  
 /organism="Arabidopsis thaliana"  
 /mol\_type="mRNA"  
 /strain="Col-0"  
 /db\_xref="taxon:3702"  
 /clone="G57FB532F04"  
 /tissue\_type="flowers and buds"  
 /plasmid="pCINSPORT\_6"  
 complement(1..927)  
 /gene="At2g39060"

## ORIGIN

Alignment Scores:  
 Pred. No.: 1.31e-68 Length: 927  
 Score: 680.50 Matches: 135  
 Percent Similarity: 73.90% Conservative: 49  
 Best Local Similarity: 54.22% Mismatches: 62  
 Query Match: 50.30% Indels: 3  
 DB: 3 Gaps: 1

US-09-743-885a-1 (1-265) x CNS0A9RJ (1-927)

```

QY 4 LeuArgAlaAspAspLeuSerPheIlePheGlyLeuLeuGlyAlaValSerPheMet 23
DB 34 CTGAGGTTATGAATAATGCTTTCTTCTGCTCTTATGCAACATGATGCTTTGGG 93
QY 24 ValPheLeuAlaProValProThrPheTyrIleTyrIleValSerSerGluGly 43
DB 94 GTGCTTGTGACACAGTGCACGCTTTTATGAGATATACAAAGAAATCATCAAGGG 153

```

```

Qy 44 TyGlnAlaIleProTyrMetValAlaLeuPheSerAlaGlyLeuLeuTyrTyrAla 63
Db 154 TTTGAGTCGATACCGTACATATGTCAGTACAGTCACTTCTCTTCTACACGGA 213
Qy 64 TyTleuArgLysAsnAlaTyrLeuIleValSerIleLeuGlyPheGlyCysAlaIleGlu 83
Db 214 ATCTGATGAGACACATGCTATCTATCATATGATATTAACACTTGTGATTCATTGAA 273
Qy 84 LeuThrTyrIleSerIlePheLeuPheTyrAlaProArgLysSerIlePheTyrGly 103
Db 274 ATCCCTACTTGTCTCTATATCTTATATCCACCAAGAGGCCAAGATTCACAGCTG 333
Qy 104 TyPheLeuLeuLeu---GluLeuGlyAlaLeuGlyMetValMetProIleThrTyrLeu 122
Db 334 AAGTTGATAGTACATCATCGTGGAGCTTGTCTCTTAATTTCTTATTAATCTT 393
Qy 123 LeuAlaGluGlySerIleArgValMetIleValGlyTyrPheCysAlaIleAsnVal 142
Db 394 TTGGTTCCAAAACACACCGGCTCGACCGTTGGATGGTTGTGCTTACAGTCTC 453
Qy 143 AlaValPheAlaIleProLeuSerIleMetArgIleValIleTyrThrLysSerValGlu 162
Db 454 GCATCTTGTGCTTCTCTTATAGCTCATGAGAGGTTAAAGACAAAGAGTGTGGA 513
Qy 163 PheMetProPheThrLeuSerLeuPheLeuThrLeuCysAlaIleThrMetTyrPheTyr 182
Db 514 TACATGCCGTTTCTTCTCTCTCTGCTCTGACTTAAAGCCGCTCATGTGTTCTTTAT 573
Qy 183 GlyPheLeuLysValPheTyrIleAlaPhe--ProAsnIleLeuGlyPheLeuPheG 202
Db 574 GGACTTCTTATCAAGACAGATGCTATGCTATGCCCAACATTCGCTTCTATTCG 633
Qy 202 TyIleValGlnMetLeuLeuTyrPheValTyrLysAspSerLysArgIleAspAspGlu 222
Db 634 GTGATGCTCATATATCTATCATATGATGATCAAGTTCACAGAAAAGCATTTGCCAA 693
Qy 222 YsSerAspProValArgGluAlaThrLysSerLysGluGlyValGluIleIleLeu 242
Db 694 CAATAAACAACCTCAAAATTAACCGATGTTAAAGTAAGTCCGATCGTGGCCGTGAAT 753
Qy 242 LeGluAspAspAsnSerAspAsn 249
Db 754 TGCTGATGTGGATCAGACAAAT 776

RESULT 7
CF445200 765 bp mRNA linear EST 04-SEP-2003
LOCUS EST681545 normalized cDNA library of onion Allium cepa cDNA clone
DEFINITION ACAG123, mRNA sequence.
ACCESSION CF445200
VERSION CF445200.1 GI:34467902
KEYWORDS EST.
SOURCE Allium cepa (onion)
ORGANISM Allium cepa
Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Asparagales; Alliaceae;
Allium.
REFERENCE 1 (bases 1 to 765)
AUTHORS Hovey M.J., Cheung F., Van Aken S., Uterback T. and Town C.D.
TITLE Expressed Sequence Tags from a normalized library of mixed onion
JOURNAL Unpublished (2003)
COMMENT Contact: Hovey MJ
Department of Horticulture
USDA-ARS and University of Wisconsin
1575 Linden Drive, Madison, WI 53706, USA
Tel: 608-262-1830
Fax: 608-262-4743
Email: mjhavey@facstaff.wisc.edu
TIGR sequence name ACAG123TR. For more information:
http://havey/lab.hort.wisc.edu
Seq primer: CAG GAA ACA GCT ATG ACC.
Location/Qualifiers

```

```

source
1. 765
/organism="Allium cepa"
/mol_type="mRNA"
/cultivar="Red Creole (bulbs), unknown (callus), Ebano &
Texas Legend (roots)"
/db_xref="taxon:4679"
/clone="ACAG123"
/tissue_type="Callus, roots, and young bulbs"
/clone_lib="normalized cDNA library of onion"
/notes="Vector: pCMVSPORT6.1-ccdb (Invitrogen); Site 1:
EcoRV (5'); Site 2: NotI (3'); Equal molar amounts of mRNA
from callus, roots, and young bulbs were combined to
synthesize the library. Normalization to enrich for
low-copy transcripts was performed by proprietary
techniques of Invitrogen."

ORIGIN
Alignment Scores:
Pred. No.: 2,25e-68 Length: 765
Score: 677.50 Matches: 127
Percent Similarity: 77.88% Conservative: 42
Best Local Similarity: 58.53% Mismatches: 47
Query Match: 50.07% Indels: 1
DB: Gaps: 1

US-09-743-885a-1 (1-265) x CF445200 (1-765)
Qy 10 SerPheIlePheGlyLeuLeuGlyAsnIleValSerPheMetValPheLeuAlaProVal 29
Db 101 GCTTTCCCTTGTGCTCTACTAGTAAGTATCTGTCTGCTCATGTGATCTTCTCCATT 160
Qy 30 ProThrPheTyrLysIleTyrLysArgLysSerSerGluGlyTyrGlnAlaIleProTyr 49
Db 161 CCAACATTTTACCGCATACACAAGAAATCAACGTAGATTTAGTGGCTTCCTTAC 220
Qy 50 MetValAlaLeuPheSerAlaGlyLeuLeuTyrTyrAlaTyrLeuArgLysAsnAla 69
Db 221 GTTGTCCTCTATTATGCTATGCTATGATATACATCAACCATTCACCAAGCAAGAG 280
Qy 70 TyIleIleValSerIleLeuGlyPheGlyCysAlaIleGluLeuThrTyrIleSerLeu 89
Db 281 TACCTTCTCATCACTATCAACAACCTTGGATGTTTCGTAACAATTTACATATGTCATG 340
Qy 90 PheLeuPheTyrAlaProArgLysSerLysIlePheThrGlyTyrLeu---MetLeuLeu 108
Db 341 TTTCTCACCTACGCCCCCAAAAAGCAAGGTGCAACCGCAAAAGATTCATCATCTCTCTA 400
Qy 109 GluLeuGlyAlaLeuGlyMetValMetProIleThrTyrTyrLeuAlaGluGlySerHis 128
Db 401 AATGTCGGGCTATTCTCTTAATTTGTTCTATCAACTCTTCTCCCTTGCCAGGGGCTAGC 460
Qy 129 ArgValMetIleValGlyTyrPheCysAlaIleAsnValAlaValPheAlaIlePro 148
Db 461 CGCCAAAACTACTTGGATGGATTTGGCTGAGCTTTCAGTCACTGATCTTTCGCTGACT 520
Qy 149 LeuSerIleMetArgIleValIleTyrThrLysSerValGluPheMetProPheThrLeu 168
Db 521 TTAAAGATTAATCAGGCTAATTAATTAAGACAAAGCGTCGATTTAGCCATTCGGCTTG 580
Qy 169 SerLeuPheLeuThrLeuCysAlaIleThrMetTyrPhePheTyrGlyIlePhePheLysValAsp 188
Db 581 TCTTTTTCCTCATCATGAGGCTATCATCTGTTCTCTTAATGATTTTCATCAACAAAGAC 640
Qy 189 PheTyrIleAlaPheProAsnIleLeuGlyPheLeuPheGlyIleValGlnMetLeuLeu 208
Db 641 ATCTATATAGGATTTCCAAATATATCTGGATTCAGCTTTGGAGTGTTCAAGATATCTTG 700
Qy 209 TyrPheValTyrLysAspSerLysArgIleAspAspGluLysSerAspPro 225
Db 701 TACATATCTCAACGATGAGAGGTTCTTTAATAGATAGCTCGAGTTAACCC 751

RESULT 8
AJ796299

```

LOCUS AJ796299 731 bp mRNA linear EST 11-AUG-2004  
 DEFINITION AJ796299 Antirrhinum majus whole plant Antirrhinum majus cDNA clone  
 018.3.10\_06, mRNA sequence.  
 ACCESSION AJ796299  
 VERSION AJ796299.1 GI:51111627  
 KEYWORDS EST.  
 ORGANISM Antirrhinum majus (snapdragon)  
 SOURCE Antirrhinum majus  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 asteridae; lamiales; Plantaginaceae; Antirrhineae;  
 Antirrhinum.  
 1 (bases 1 to 731)  
 Ziegler, J., Stuber, K., Saedler, H., Sommer, H. and Schwarz-Sommer, Z.  
 Antirrhinum EST collection  
 Unpublished (2003)  
 CONTACT: Schwarz-Sommer Z  
 MOLEKULARE PFLANZENGENETIK  
 MPI fuer Zuechtungsforschung  
 Carl-von-Linne Weg 10, D-50829, Germany.  
 FEATURES  
 source  
 1. 731  
 /organism="Antirrhinum majus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:4151"  
 /clone="018.3.10\_06"  
 /tissue\_type="whole plant"  
 /clone\_lib="Antirrhinum majus whole plant"  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 2 38e-67 length: 731  
 Score: 668.50 Matches: 123  
 Percent Similarity: 78.97% Conservative: 46  
 Best Local Similarity: 57.48% Mismatches: 43  
 Query Match: 49.41% Indels: 2  
 DB: 1 Gaps: 1  
 US-09-743-885a-1 (1-265) x AJ796299 (1-731)  
 QY 6 AAlaAspAspLeuSerPheIlePheGlyLeuLeuGlyAValIleValSerPheMetValPhe 25  
 Db 86 AGTAACCAATGGCTTTCGGCTTTCGGCTTTCAGGCAACATTTGCTTCATGCTTTTC 145  
 QY 26 LeuAlaProVal-ProThrPheTyrllyIleTyrllyAsnGlySerSerGluGlyTyrl 45  
 Db 146 CTGGACCCAGTNGCCGACCTTTATATAAATATACAGAGAAATCAAGTCAAGGGTTCCA 205  
 QY 45 nAlaIleProTyMetValAlaLeuPheSerAlaGlyLeuLeuLeuTyrllyAlaTyrl 65  
 Db 206 AAGTGAACCTTAATGCTGGGCTTATTCAGTCCAGCTGCGATATCTATGATTTCT 265  
 QY 65 uATGlyAsnAlaTyrlleuIleValSerIleAsnGlyPheGlyCysAlaIleGluLeuTh 85  
 Db 266 CAAAGCAGACACACCTTCTCATTCATTAATCACTGCGATGTTTTCAAACCAT 325  
 QY 85 rTyrlleSerLeuPheLeuPheTyrlaProAlaGlySerIlePheThrGlyTyrl 105  
 Db 326 ATACATCTGCTTTATTTTATACGCAACAAAAACACAAAGTCAAACTATGAAGCT 385  
 QY 105 uMet---LeuLeuGluLeuGlyAlaLeuGlyMetValMetProIleThrTyrlleuAl 124  
 Db 386 TATTTTGTCTAAACATCTTGGGATGGGATGATGCTGCTAATTTATTTAGC 445  
 QY 124 agGlySerIleArgValMetIleValGlyTyrlleCysAlaAlaIleAsnValAla 144  
 Db 446 CAAGGCTCAACCCGTCACAAACATTTGGGTGATTTGCTGATATTCCTTGCTGTGT 505  
 QY 144 lPheAlaAlaProLeuSerIleMetArgIleValIleTyrllySerValGluPhe 164  
 Db 506 ATTATATGCCCCCTTAAAGCTGGAGACAAAGTTATCGCATTAAGAGTGTAGAGTTAT 565  
 QY 164 cProPheThrLeuSerLeuPheLeuThrLeuCysAlaIleThrMetIlePhePheTyrl 184

Db 566 GCCATTTCCTTGTGCTTTTCTTAACCTAAGTGTGATATGATGCTTTCACGGCTT 625  
 QY 184 ePheIleValAspPheTyrlleAlaPheProAsnIleLeuGlyPheLeuPheTyrlleVal 204  
 Db 626 ATTCGCCAAAGACTCAAAATGCTTATTCCTTAATGCTTGGGTGATCTTTGAGTGT 685  
 QY 204 lGluMetLeuLeuTyrlPheValTyrllyAspSerIleValArg 217  
 Db 686 TCAAATGATATCTTATGATATATTAATTAAGACGCAAGAA 725  
 RESULT 9  
 BU011719 665 bp mRNA linear EST 22-AUG-2002  
 LOCUS BU011719  
 DEFINITION OGJ16P24.yg ab1 OG\_EFGHU lettuce serriola lactuca sativa cDNA clone  
 OGJ16P24, mRNA sequence.  
 ACCESSION BU011719  
 VERSION BU011719.1 GI:22446114  
 KEYWORDS EST.  
 SOURCE Lactuca sativa  
 ORGANISM Lactuca sativa  
 Lactuca sativa  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 asteridae; campanulids; Asterales; Asteraceae; Cichorioideae;  
 Cichorieae; Lactuca.  
 1 (bases 1 to 665)  
 Kozik, A., Michemore, R.W., Knapp, S., Matvienko, M., Rieseberg, L.,  
 Lin, H., van Damme, M., Lavelle, D., Chevalier, P., Ziegler, J.,  
 Ellison, P., Kolman, D., Slabaugh, M.S., Livingston, K., Zhou, Y.,  
 Lai, Z., Church, S., Jackson, L. and Bradford, K.  
 Lettuce and Sunflower ESTs from the Composite Genome Project  
 http://comgenomics.ucdavis.edu/  
 unpublished (2002)  
 CONTACT: Alexander Kozik [R.W.Michemore]  
 Department of Vegetable Crops, R.W.Michemore Lab  
 University of California at Davis (UCD)  
 Aamundson Hall, UCD, Davis, CA 95616, USA  
 Tel: 1-(530)-742-1742  
 Fax: 1-(530)-752-9659  
 Email: akozik@ucdavis.edu [michemore@vegmail.ucdavis.edu]  
 belongs to config OG\_CA.Config93942, see http://cgpdb.ucdavis.edu/  
 for details.  
 Plate: OGJ16 row: P column: 24.  
 FEATURES  
 source  
 1. 665  
 /organism="Lactuca sativa"  
 /mol\_type="mRNA"  
 /cultivar="L.serriola"  
 /db\_xref="taxon:4236"  
 /clone="OGJ16P24"  
 /lab\_host="E.coli"  
 /clone\_lib="OG\_EFGHU lettuce serriola"  
 /note="Vector: pBRCDNA5f1AB; The library was constructed  
 from 10 different sources of RNA from a single genotype.  
 Separate cDNAs were generated using primers that  
 incorporated unique 5' and 3' tags to distinguish each  
 source of RNA. cDNAs were then pooled, size-fractionated,  
 directionally cloned into a custom medium-copy vector and  
 transformationally made with four size classes to minimize  
 size bias. Details of each source of RNA and library  
 construction can be obtained at http://cgpdb.ucdavis.edu/  
 TAG\_TISSUE=flowers pre-fertilized  
 TAG\_LIB=OG\_EFGHU lettuce serriola  
 TAG\_SEQ=CGTTGACGGG"  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 1 8e-66 length: 665  
 Score: 660.50 Matches: 122  
 Percent Similarity: 75.35% Conservative: 40  
 Best Local Similarity: 56.74% Mismatches: 52  
 Query Match: 48.82% Indels: 1  
 DB: 5 Gaps: 1



Db	768	GAACAATGAAAGATCAATGATGCTCCTTACCACTTAGTGAGCCCTGAAGTACCC	822
Qy	238	l e l e l e a n l e g l u a s p h e n s e r ----- a s p a n	249
Db	828	GTAGTTTGATATTACACCTGATGATCAAGAAGAAATAT	872
RESULT 11			
LOCUS	CO123087		
DEFINITION	GR_EB05A21.f GR_Eb Gossypium raimondii cDNA clone GR_EB05A21 5'		
ACCESSION	CO123087		
VERSION	CO123087.1		
KEYWORDS	EST.		
SOURCE	Gossypium raimondii		
ORGANISM	Gossypium raimondii		
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosidae; eutroside II; Malvales; Malvaceae; Malvaceae; Gossypium.		
AUTHORS	1 (bases 1 to 861) Udall,J.A., Rapp,R.A., Wendel,J.F., Rao,K., Soderlund,C. and Ming,R.A.		
TITLE	Global assembly of Cotton ESTs		
JOURNAL	Unpublished (2004)		
COMMENT	Contact: Rod A. Ming Arizona Genomics Institute The University of Arizona Forbes Building Room 303, Tucson, AZ, 85721-0036, USA Tel: 520 626 9595 Fax: 520 621 1259 Email: http://genome.arizona.edu Plate: 05 row: A column: 21.		
FEATURES	Location/Qualifiers		
Source	1..861 /organism="Gossypium raimondii" /mol_type="mRNA" /db_xref="taxon:29730" /clone="GR_EB05A21" /tissue_type="floral" /dev_stage="3 to +3 DPA" /lab_host="NDH10B" /clone_lib="GR_Eb" /note="vector: pCMV.SPORT-6.1; Site 1: NotI; Site 2: EcoRV; library made by Invitrogen with RNA supplied by Wendle lab. Directional cloned into NotI-EV. Colonies plated/picked by AGI. More glycerol clones held in -80."		
ORIGIN			
Alignment Scores:			
Pred. NO.:	8,866-66	Length:	861
Score:	656.00	Matches:	126
Percent Similarity:	70.59%	Conservative:	54
Best Local Similarity:	49.41%	Mismatches:	61
Query Match:	48.48%	Indels:	14
DB:	7	Gaps:	3
US-09-743-885A-1 (1-265) x CO123087 (1-861)			
Qy	9	LeuserPheillePhcGlyLeuLeuGlyYanIlleValSerPheMetValPheIleuAlaPro	28
Db	112	CTGGCTGTGTGTTTGGTATATATTAGTAAACATCTCAGTCTCTATTCTTGGGACCA	171
Qy	29	ValProThrPheTyLysIleTyrTyLysArgLysSerSerGluGlyTyGlnAlaIlePro	48
Db	172	GTGCCAACATTTCTGCGGATTACAAAAGAAATGACATGAGAGTTTCCAGTCACCA	231
Qy	49	TyMetValAlaLeuPheSerZalaglyLeuLeuLeuTyTyTyAlaTyLeuArgLysAsn	68
Db	232	TACCAAGTTGATTCTTTCAGTTCATGCATGCATAGGCTGTATTACGCGTTGATTAAGG	291
Qy	69	AlaTyLeuIleValSerIleAnGlyPheGlyCYsAlaIleGluLeuThrTyTyrIleSer	88

Db	Accession	Source	Organism	Reference Authors	Journal Comment	Features
Db	292	GCTTCCCTTCGATACACATCAACGCTTTTGGATGTGTTGGACATATTATATATCT	351			
Qy	89	LeupheteuPheTyrrAlaProArgLysSerIysAlaIlePheThrArgLTyPLeuMetLeu---	107			
Db	352	ATGTTCTTGGCTTATGCTTCMAAAGATATGAGATGCGGCTATGAACTTTTATATCA	411			
Qy	108	LeuGIleuLeuGIyAlaLeuGIyMetValMetProIleThrTyrrLeuLeuAlaGIuGIySer	127			
Db	412	ATGAACCTGGGGCTCTTCTCTTTATTTCTACCTCCACACCTTTCTGCTGAAAAGCTCG	471			
Qy	128	HisArgValMetIleValGIyTrrPLeuCysAlaAlaIleAsnValAlaValPheAlaAla	147			
Db	472	ATTCGGCGCCAAAGTTCTTGGCTGATTTGTGTGCCATCTCTGATCCGCTTTTGACAGA	531			
Qy	148	ProLeuSerIleMetArgInValIleIleLysThrLysSerValGIuPheMetProPheThr	167			
Db	532	CCCTTAAACATTATGAGCAGATATACGAACAAAGCGTTGATCATCTCCCTTTACCC	591			
Qy	168	LeuSerLeuPheLeuThrLeuCysAlaThrMetTrrPhePheTyrrGIyPhePheLys	187			
Db	592	TTATCATTTTTCCTCACATTCAGTCAGCATATATGATTTGCTTATGACATTTTATTAAG	651			
Qy	188	AspPheTyrrIleAlaPheProAsnIleLeuGIyPheLeuPheGIyIleValGIuMetLeu	207			
Db	652	GACCTCTGTAGCTCTCCCTAAATGATCTAGCTTTCATCTTGGGGATGCTCCAGATGCTT	711			
Qy	208	LeuTyrrPheValTyrrLysAspSerLysArgIleAspArgGIuLysSerAspProValArg	227			
Db	712	CTTACAGCCATTATACAGACAGTGAAGGTTAACTATAGAGAAAGAAACCTACCACT	771			
Qy	228	GIuAlaThrLysSer-----LysGIuGIyValGIu	237			
Db	772	GAAACAATGAAACATCATCATGCTCTTACCACATTAATGCTCTGAAGTTACACCC	831			
Qy	238	IleIleIleAsnIleGIuAspAspAspSer-----AspAsn	249			
Db	832	GTAAGTTTGATTCATTCACATCGATGATACAAAGAGAAGATATAT	876			
RESULT 12						
LOCUS	CO125993					
DEFINITION	GR_EB09M10.f GR_EB Gossypium raimondii cDNA clone GR_EB09M10 5',					
ACCESSION	CO125993					
VERSION	CO125993.1					
KEYWORDS	EST.					
SOURCE	Gossypium raimondii					
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;					
REFERENCE	Kim,H.,Yu,Y.,Kudrna,D.,Hatfield,J.,Stum,D.,Weller,C.,					
AUTHORS	Udall,J.A.,Rapp,R.A.,Wendel,J.F.,Rao,K.,Soderlund,C. and					
TITLE	Wing,R.A.					
JOURNAL	Global assembly of Cotton ESTs					
COMMENT	Unpublished (2004)					
	Contact: Rod A. Wing					
	Arizona Genomics Institute					
	The University of Arizona					
	Forbes Building Room 303, Tucson, AZ, 85721-0036, USA					
	Tel: 520 626 9595					
	Fax: 520 621 1259					
	Email: http://genome.arizona.edu					
	Plate: 09 row: M column: 10.					
FEATURES						
Source						
	1..865					
	/organism="Gossypium raimondii"					
	/mol_type="mRNA"					
	/db_xref="taxon:29730"					
	/clone="GR_EB09M10"					
	/tissue_type="floral"					
	/dev_stage="_3 to +3 DPA"					

/lab\_host="DH10B"  
/clone\_id="GR\_Eb"  
/note="Vector: pCMV.SPORT-6.1; Site 1: NotI; Site 2:  
ESORV; Library made by Invitrogen with RNA supplied by  
Wendle lab. Directional cloned into NotI-EV. Colonies  
plated/picked by AGI. More glycerol clones held in -80."

## ORIGIN

## Alignment Scores:

Pred. No.: 9,888-66 Length: 865  
Score: 655.50 Matches: 124  
Percent Similarity: 71.20% Conservative: 54  
Best Local Similarity: 49.60% Mismatches: 61  
Query Match: 48.45% Indels: 11  
DB: 7 Gaps: 2

US-09-743-885A-1 (1-265) x CO125993 (1-865)

QY 9 LeuSerPheIlePheGlyLeuLeuGlyValAsnIleValSerPheMetValPheLeuAlaPro 28  
Db 114 CTGGCTGTTGTTTGGTATATAGTAAACATCATCTCAGCTCGTATTTCTGGACACA 173  
QY 29 ValProThrPheTyrlsIleTyrlsArglySerSerGluGlyTyrlsGlnAlaIlePro 48  
Db 174 GTGCACAACATTCCTCGATTACAAAAGAAATCGACTGAGAGTTTCAGTCCAGTCCCA 233  
QY 49 TyrMetValAlaLeuPheSerAlaGlyLeuLeuLeuTyrlsTyrlsAlaTyrlsLeuArglyAsn 68  
Db 234 TACACAGTTGATTTGATTCAGTTCATAGCGCTGATTCACGCTGATTAATAAAGGT 293  
QY 69 AlaTyrlsLeuIleValSerIleAsnGlyPheGlyCysAlaIleGluLeuThrTyrlsSer 88  
Db 294 GCTTCCCTTCGATCACCATCAACGCTTTGATGTGTGTGAGACTATTATATATCT 353  
QY 89 LeuPheLeuPheTyrlsAlaProArglySerSerIlePheThrGlyTyrlsLeuMetLeu--- 107  
Db 354 ATGTTCTGCTTATGCTTCAAGAAATGACAGATGCGCTATGAACCTTTTATATCA 413  
QY 108 LeuGluLeuGlyAlaLeuGlyMetValMetProIleThrTyrlsLeuAlaGluGlySer 127  
Db 414 ATGAACCTGGGGCTTCTCTTCATCTCATCTCCACACCTTTCCTGCAAAAGCTCG 473  
QY 128 HisArgValMetIleValGlyTyrlsIleCysAlaIleAsnValAlaValPheAlaIle 147  
Db 474 ATTCGCGTCCAAAGTCTTGCTGCTGATTTGCTCCATCTCTGATCCGCTTTGACACA 533  
QY 148 ProLeuSerIleMetArgGlnValIleTyrlsPheTyrlsSerValGluPheMetProPheThr 167  
Db 534 CCTTAACATTAATGACACGAGTTATGCAACAAAGACGTTGAGTTCAATGCTTTCCACC 593  
QY 168 LeuSerLeuPheLeuThrIleCysAlaIleThrMetTyrlsPhePheTyrlsGlyPhePheIleLys 187  
Db 594 TTATCATTTTTCCTCACATTAAGTACAGTATGTTGTTGCTTATGACATTTATTAAG 653  
QY 188 AspPheTyrlsIleAlaPheProAsnIleLeuGlyPheLeuPheGlyIleValGlnMetLeu 207  
Db 654 GACCTCTGTGATGCTCTCCCTAATGATCTAGCTTCACTTGGGAGATCTCCAGATGCTT 713  
QY 208 LeuTyrlsPheValTyrlsAspSerLysArgIleAspArgLysSerAspProValArg 227  
Db 714 CTCTACCCCATTTACACACAGTGAAGGTTAACATAGAGAAAGAAAGACTCCAGCT 773  
QY 228 GluAlaThrLysSer-----LysGluGlyValGlu 237  
Db 774 GAACAATAAGAAACATCAATGCTGCTCTTACCATTAAGTGCCTCGAAAGTTCACCCC 833  
QY 238 IleIleIleAsnIleGluAspAspAsnSer 247  
Db 834 GTAGTTTGGATTTACACATGATGATACA 863

RESULT 13  
CO121875  
LOCUS

CO121875 888 bp mRNA linear EST 16-JUN-2004

DEFINITION GR\_Eb03E02.f GR\_Eb Gossypium raimondii cDNA clone GR\_Eb03E02 5',  
mRNA sequence.  
ACCESSION CO121875  
VERSION CO121875.1 GI:48820562  
KEYWORDS EST.  
SOURCE Gossypium raimondii  
ORGANISM Gossypium raimondii  
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.  
1 (bases 1 to 888)  
Kim,H., Yu,Y., Kudrna,D., Hatfield,J., Stum,P., Mueller,C.,  
Udall,J.A., Rapp,R.A., Wendel,J.F., Rao,K., Soderlund,C. and  
Wing,R.A.  
Global assembly of Cotton ESTs  
Unpublished (2004)  
Contact: Rod A. Wing  
Arizona Genomics Institute  
The University of Arizona  
Forbes Building Room 303, Tucson, AZ, 85721-0036, USA  
Tel: 520 626 9595  
Fax: 520 621 1259  
Email: <http://genome.arizona.edu>  
Plate: 03 row: E column: 02.  
Location/Qualifiers

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

## FEATURES

## source

1..888  
/organism="Gossypium raimondii"  
/mol\_type="mRNA"  
/db\_xref="taxon:29730"  
/clone="GR\_Eb03E02"  
/tissue\_type="floral"  
/dev\_stage="3 to +3 DPA"  
/lab\_host="DH10B"  
/clone\_id="GR\_Eb"  
/note="Vector: pCMV.SPORT-6.1; Site 1: NotI; Site 2:  
ESORV; Library made by Invitrogen with RNA supplied by  
Wendle lab. Directional cloned into NotI-EV. Colonies  
plated/picked by AGI. More glycerol clones held in -80."

## ORIGIN

## Alignment Scores:

Pred. No.: 1,028-65 Length: 888  
Score: 655.50 Matches: 124  
Percent Similarity: 71.20% Conservative: 54  
Best Local Similarity: 49.60% Mismatches: 61  
Query Match: 48.45% Indels: 11  
DB: 7 Gaps: 2

US-09-743-885A-1 (1-265) x CO121875 (1-888)

QY 9 LeuSerPheIlePheGlyLeuLeuGlyValAsnIleValSerPheMetValPheLeuAlaPro 28  
Db 118 CTGGCTGTTGTTTGGTATATAGTAAACATCATCTCAGCTCGTATTTCTGGACACA 177  
QY 29 ValProThrPheTyrlsIleTyrlsArglySerSerGluGlyTyrlsGlnAlaIlePro 48  
Db 178 GTGCACAACATTCCTCGATTACAAAAGAAATCGACTGAGAGTTTCAGTCCAGTCCCA 237  
QY 49 TyrMetValAlaLeuPheSerAlaGlyLeuLeuLeuTyrlsTyrlsAlaTyrlsLeuArglyAsn 68  
Db 238 TACACAGTTGATTTGATTCAGTTCATAGCGCTGATTCACGCTGATTAATAAAGGT 297  
QY 69 AlaTyrlsLeuIleValSerIleAsnGlyPheGlyCysAlaIleGluLeuThrTyrlsSer 88  
Db 298 GCTTCCCTTCGATCACCACCAACGCTTTGATGTGTGTGAGACTATTATATATCT 357  
QY 89 LeuPheLeuPheTyrlsAlaProArglySerSerLysIlePheThrGlyTyrlsLeuMetLeu--- 107  
Db 358 ATGTTCTGCTTATGCTTCAAGAAATGACAGATGCGCTATGAACCTTTTATATCA 417  
QY 108 LeuGluLeuGlyAlaLeuGlyMetValMetProIleThrTyrlsLeuAlaGluGlySer 127  
Db 418 ATGAACCTGGGGCTTCTCTTCATCTCATCTCCACACATTTCTGCTGAAGGCTCG 477

Qy		128	HlsarqValmctiIevaIGlYTriplcCyslaalaiaileanValAlaValPhealAala	147
Dd		478	ATTGGCGTCGAAGTCTTGCGCTGCATTTGTGTCGCATCTCTGTATCCGCTCTTTCACACA	537
Qy		148	ProlenSerilImeLarGlnValilleYsrThrlvSsrValGlunPhewEProberThr	167
Dd		538	CCCTTAACAATTATGACACAGATGTATACCAACAAGAAGCGTTGAATCATGCCCTTTCACC	597
Qy		168	LenseurLeupheleuthrLeucCyalaAhmrnecTrpphepheTygLyphetheylslys	187
Dd		598	TTATCATTTTCCCACATTCATGATGACATAATGCGTTTGCTTAAGACTAATTTATTAAG	657
Qy		188	AapPheTYrlleaIaphPeProAsmlleuGlyPheleuPheglYilaIGlmnetLeu	207
Dd		658	GACCTCTGTGTAGCTCTCCCTAACTACTTAGGCTTCATCTTGGGAGATGCTCCAAGTCTT	717
Qy		208	LeutryPhevalTYrlvAspSerlylarGllleaPaSpGluLvSsrAspBovalArg	227
Dd		718	CTTACGCGCATTTACAGACACAGTGAAGAAGTTAACATAGAGGAAGAAGCTACAGCT	777
Qy		228	GLuAlarHlrYsser-----LYgLIgLYlaIGlu	237
Dd		778	GAACAATAATGAAGAATCATCATATGTCGCTCTTACCATTAGTGCTCTGAAGTTACACCC	837
Qy		238	IlellellealeenilleGluAspAspaasr	247
Dd		838	CTACTTTTGATATTCACACTGATGATACCA	867
RESULT 14				
LOCUS		COL121803		
DEFINITION		COL121803	mRNA	EST 16-JUN-2004
			gb EB03C11.f_GR_EB_Goseypium_raimondii cDNA clone GR_EB03C11 5'	
ACCESSION		COL121803		
VERSION		COL121803.1		
KEYWORDS		GI:48820490		
SOURCE		EST.		
ORGANISM		Goseypium raimondii		
		Goseypium raimondii		
		Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
		Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;		
		rosidae; eurosoids II; Malvales; Malvaceae; Malvoideae; Gossypium.		
REFERENCE		1 (bases 1 to 865)		
AUTHORS		Kim,H., Yu,Y., Kudrna,D., Hatfield,J., Stum,D., Mueller,C.,		
		Udall,J.A., Rapp,R.A., Wendel,J.F., Rao,K., Soderlund,C. and		
		Wing,R.A.		
TITLE		Global assembly of Cotton ESTs		
JOURNAL		Unpublished (2004)		
COMMENT		Contact: Rod A. Wing		
		Arizona Genomics Institute		
		The University of Arizona		
		Forbes Building Room 303, Tucson, AZ, 85721-0036, USA		
		Tel: 520 626 9595		
		Fax: 520 621 1259		
		Email: http://genome.arizona.edu		
FEATURES		plate: 03 row: C column: 11.		
source		location/Qualifiers		
		1..865		
		/organism="Goseypium raimondii"		
		/mol_type="mRNA"		
		/db_xref="taxon:29730"		
		/clone="GR_EB03C11"		
		/tissue_type="floral"		
		/dev_stage="3 to +3 DPA"		
		/lab_host="DH10B"		
		/clone_id="GR_EB"		
		/note="vector: pCMV.SPORT-6.1; site_1: NotI; site_2:		
		NotcIV; library made by Invitrogen with RNA supplied by		
		Wendel lab. Directional cloned into NotI-EV. Colonies		
		plated/picked by AGI. More glycerol clones held in -80."		
ALIGNMENT Scores:				
ORIGIN				

Pred. No.:	1,296-65	length:	865
Score:	654.50	Matches:	124
Percent Similarity:	71.08%	Conservative:	53
Best Local Similarity:	49.80%	Mismatches:	61
Query Match:	48.37%	Indels:	11
DB:	7	Gaps:	2

  

US-09-743-685A-1 (1-265) x CO121803 (1-865)	
Qy 9 LeuSerPheIlePheGlyLeuGlyLeuGlyValAsnIleValSerPheMetValPheLeuAlaPro 28	119 CTGCGTGTGTTTGGTATATATAGTAACATCATCTACGCTCCTGATATTTCTGGACACA 178
Qy 29 ValProThrPheTyrlsIleTyrlsArglySerSerGergluglyTyrlGlnAlaIlePro 48	179 GTGGCAACATTCCTCGAATTTACAAAAGAAATCGACTGAGATTTCCATCCTGCA 238
Qy 49 TyrMetValAlaIleuPheSerAlaGlyLeuLeuLeuTyrlTyrlAlaTyrlLeuArglyLeuAsn 68	239 TACACAGTTGCATTTGTTCACTGATGCATGCTATGCTGATTTACGGCTTGATTAAGGT 298
Qy 69 AlaTyrlLeuIleValSerIleAsnGlyPheGlyCysAlaIleGluLeuThrTyrlIleSer 88	299 GCTTTCCTTCGATCACCATCAACGCTTTTGATGCTGATGATGAGACATATATATATCT 358
Qy 89 LeuPheLeuPheTyrlAlaProArglySerIleSerIlePheThrGlyTyrlPheMetLeu--- 107	359 ATGTTCTTGGCTTATGCTTCAAAAGAAATGCGATGCGCTATGAAACCTTTATATATCA 418
Qy 108 LeuGluLeuGlyAlaIleuGlyMetValMetProIleThrTyrlLeuLeuAlaGluGlySer 127	419 ATGAACCTGGGGCTCTTCTCTTTCATTTCTACCTCAACACTTTGCTGGAAGAAGCTCG 478
Qy 128 HisArgValMetIleValGlyTyrlPheCysAlaIleAlaIleAsnValAlaValPheAlaAla 147	479 ATTCGGCTCCAAAGTTCTTGGCTGATTTGTTGTTGCCATCTCTGATTCGGTCTTTGGACGA 538
Qy 148 ProLeuSerIleMetArgIleValIleIleTyrlsSerValGluPheMetProPheThr 167	539 CCCTTAAACATTATGACACAGGATTAACGAACAAAGACGTTGATTCATCCTTTCCACC 598
Qy 168 LeuSerLeuPheLeuThrLeuCysAlaIleThrMetTrpPhePheTyrlGlyPhePheValys 187	599 TTATCATTTTTCCTCAACATTTGATGACGATATGATGTTGCTTATGACTATTTATTAAG 658
Qy 188 AspPheTyrlIleAlaPheProAsnIleLeuGlyPheLeuPheGlyIleValGlnMetLeu 207	659 GACCTCTGTGTAGCTCTCCCTTAATGTAAGTCTTCACTTTGGGAGATGCTCAATGCTT 718
Qy 208 LeuTyrlPheValTyrlsAspSerIleAspArgIleAspAspGluIysSerAspProValArg 227	719 CTTACGCGCATTTACAGACACAGGAAAGGTTAATACATAGGAAAGAAAGCTACAGCT 778
Qy 228 GluAlaThrIysSer-----LysGluGlyValGlu 237	
Qy 779 GAAACAAATGAAACCATCATATGCTGCTCTTACCATATGATGCTCTGAAGTTACACCCC 838	
Qy 238 IleIleIleAsnIleGluAspAspAsn 246	
Qy 839 GTAGTTTGGATTTACACATGATGAT 865	

  

RESULT 15	
LOCUS CO090096	858 bp mRNA linear EST 16-JUN-2004
DEFINITION GR_Ea09005.f GR_Ea Gossypium raimondii cDNA clone GR_Ba09005 5',	
ACCESSION CO090096	
VERSION CO090096.1	GI:48784654
KEYWORDS EST.	
SOURCE Gossypium raimondii	
ORGANISM Gossypium raimondii	
Spmetophyta: Magnoliophyta: eudicotyledons, core eudicots:	

REFERENCE  
1 (bases 1 to 858)  
AUTHORS  
Kim, H., Yu, Y., Kudrna, D., Hatfield, J., Stum, D., Mueller, C.,  
Udall, J. A., Rapp, R. A., Wendel, J. F., Rao, K., Soderlund, C. and  
Wang, R. A.  
TITLE  
Global assembly of cotton ESTs  
JOURNAL  
Unpublished (2004)  
COMMENT  
Contact: Rod A. Wing  
Arizona Genomics Institute  
The University of Arizona  
Forbes Building Room 303, Tucson, AZ, 85721-0036, USA  
Tel: 520 626 9395  
Fax: 520 621 1259  
Email: <http://genome.arizona.edu>  
Plate: 09 row: O column: 05.

FEATURES  
source  
1. 858  
location/Qualifiers  
/organism="Gossypium raimondii"  
/mol\_type="mRNA"  
/db\_xref="taxon:29730"  
/clone="GR\_Ea09005"  
/issue\_type="whole seedlings"  
/dev\_stage="first true leaves"  
/lab\_host="DH10B"  
/clone\_id="GR\_Ea"  
/note="Vector: pCMV.SPORT-6.1; Site 1: NotI; site 2:  
EcoRV; library made by Invitrogen with RNA supplied by  
Wendle lab. Directional cloned into NotI-EV. Colonies  
plated/picked by AGI. More glycerol clones held in -80."

## ORIGIN

## Alignment Scores:

Pred. No.:	1.67e-65	Length:	858
Score:	653.50	Matches:	124
Percent Similarity:	70.97%	Conservative:	52
Best Local Similarity:	50.00%	Mismatches:	61
Query Match:	48.30%	Indels:	11
DB:	7	Gaps:	2

US-09-743-885A-1 (1-265) x CO090096 (1-858)

```

QY      9 LeuSerPheIlePheGlyLeuLeuGlyAsnIleValSerPheMetValPheLeuAlaPro 28
      115 CTGGCTGTGTTGTTGGTATGTAAGTAACATCACTCAGCTCGATATTCGACCA 174
QY      29 ValProThrPheTyrlsIleTyrlsArgLysSerSerGluGlyTyrlsAlaIlePro 48
      175 GTGCCAACATTCCTCGGATTTACAAAGAAATCGACTGAGAGTTTCAGTCCGCA 234
QY      49 TyMetValAlaLeuPheSerAlaGlyLeuLeuLeuTyrlsAlaTyrlsArgLysAsn 68
      235 TACCAAGGTGATGTTCACTGATGCTATGCTGATTAACGCGTTCATTAAGGCT 294
QY      69 AlaTyrlsLeuIleValSerIleAsnGlyPheGlyCysAlaIleGluLeuThrTyrlsSer 88
      295 GCCTTCCTTCGATCACCATCAACGCTTTGGATGTGTTGGAGACTAATTATATATCT 354
QY      89 LeuPheLeuPheTyrlsAlaProArgLysSerLysIlePheThrGlyTyrlsLeuMetLeu-- 107
      355 ATGTTCTTGGCTTATGCTTCAAGAAATGACGAGATGCGGCTATGAACCTTTTATATCA 414
QY      108 LeuGluLeuGlyAlaLeuGlyMetValMetProIleThrTyrlsLeuAlaGluGlySer 127
      415 ATGAACCTGGGGCTCTCTCTTCATTCATCTCATCCACACACTTTCGCTGAAGAGCTCG 474
QY      128 HisArgValMetIleValGlyTyrlsCysAlaAlaIleAsnValAlaValPheAlaAla 147
      475 ATTCGCGTCCAAAGTCTTGGCTGATTTGTGTCATCTGTCATCTGCTTCGTCAGCA 534
QY      148 ProLeuSerIleMetArgGlnValIleValThrLysSerValGluPheMetProPheThr 167
      535 CCTTAACATTATGACGACGAGTATACGAACAAAGAGCGTTGAGTTTCATGCTTCACCC 594

```

```

QY      168 LeuSerLeuPheLeuThrLeuCysAlaThrMetTrpPhePheTyrlsIlePheIleLys 187
      595 TTATCATTTTTCCTCCATGAGTGCAGTATGTTGCTTATGACATTAATTATAG 654
QY      188 AspPheTyrlsIleAlaPheProAsnIleLeuGlyPheLeuPheGlyIleValGluMetLeu 207
      655 GACCTCTGTAGCTCTCCCTAATGTAAGCTTCACTTGGGAGATGCTCCAGATCCTT 714
QY      208 LeuTyrlsPheValTyrlsAspSerLysArgIleAspAspGluLysSerAspProValArg 227
      715 CTCTAGCCCATTTTACAGACACAGTGAAGGTTAACAATAGAGAAAGAGCTACAGCT 774
QY      228 GluAlaThrLysSer-----LysGluGlyValGlu 237
      775 GAACAATGAAGAAACATCAATGTCGTCCTTACACATTANGTGCCTCGAAGTTACACCCC 834
QY      238 IleIleIleAsnIleGluAspAsp 245
      835 GTAGTTTGGATATTCACACTGAT 858

```

Search completed: December 24, 2004, 23:26:17  
Job time : 3873 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: December 24, 2004, 20:54:22 ; Search time 638 seconds

(without alignments)  
2312.807 Million cell updates/sec

Title: US-09-743-885A-1

Perfect score: 1353

Sequence: 1 MAQLRADDSFIFGLGNTV.....NSDNALQSMKXPSRLRTSK 265

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 4105333 segs, 2784095677 residues

Total number of hits satisfying chosen parameters: 8210666

Minimum DB seg length: 0  
Maximum DB seg length: 200000000

Post-Processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODE=frame+ p2n.model -DBV=xlp  
-Q=/cgn2\_1/USPTO.epool\_p/US09743865/runat\_23122004\_165301\_9669/app.query.fasta\_1.455  
-DB=Published Applications NA -QFMT=fastag -SUFFIX=trmpb -MINMATCH=0.1  
-LOOPEXT=0 -LOOPEXT=0 -UNITS=bites -START=1 -END=1 -MATRIX=blosum62  
-TRANS=human40.cdi -LIST=45 -DOCCALIGN=200 -THR SCORE=ext -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0  
-MAXLEN=200000000 -USER=US09743865 @CGN\_1\_1\_912 @runat\_23122004\_165301\_9669  
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOC=100  
-LONGLOG -DEV TIMEOUT=120 -MARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FAPOP=6 -FAPEXT=7 -YGAPOP=10 -YGAEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications NA: \*  
1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq: \*  
2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq: \*  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq: \*  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq: \*  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq: \*  
6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq: \*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq: \*  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq: \*  
9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq: \*  
10: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq: \*  
11: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq: \*  
12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq: \*  
13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq: \*  
14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq: \*  
15: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq: \*  
16: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq: \*  
17: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq: \*  
18: /cgn2\_6/ptodata/1/pubpna/US10F\_PUBCOMB.seq: \*  
19: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq: \*  
20: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUBCOMB.seq: \*  
21: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUBCOMB.seq: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	681	50.3	1476	16	US-10-424-599-85620 Sequence 85620, A
2	669.5	49.5	945	18	US-10-425-115-28611 Sequence 29611, A
3	654.5	48.4	1849	16	US-10-424-599-77 Sequence 77, Appl
4	654	48.3	1102	16	US-10-425-114-32824 Sequence 32824, A
5	653.5	48.3	1500	16	US-10-425-114-15869 Sequence 15869, A
6	653.5	48.3	2436	18	US-10-425-115-28621 Sequence 28621, A
7	653	48.3	1306	16	US-10-425-114-3587 Sequence 3587, A
8	653	48.3	1415	18	US-10-425-115-28623 Sequence 28623, A
9	650.5	48.1	1703	18	US-10-425-115-29613 Sequence 29613, A
10	647.5	47.9	1513	17	US-10-767-701-14622 Sequence 14622, A
11	646.5	47.6	1141	17	US-10-437-963-15131 Sequence 15131, A
12	646.5	47.6	1141	17	US-10-739-930-5395 Sequence 5395, A
13	643	47.5	1220	16	US-10-425-114-14414 Sequence 14414, A
14	643	47.5	1360	18	US-10-425-115-28616 Sequence 28616, A
15	640	47.3	1278	16	US-10-425-114-27663 Sequence 27663, A
16	638	47.2	1306	16	US-10-424-599-66 Sequence 66, Appl
17	634	46.9	1557	17	US-10-424-599-66357 Sequence 66357, A
18	633.5	46.8	1759	17	US-10-767-701-14763 Sequence 14763, A
19	629.5	46.5	1361	16	US-10-425-114-2269 Sequence 2269, A
20	629.5	46.5	1766	18	US-10-425-115-77019 Sequence 77019, A
21	623.5	46.1	879	9	US-09-938-842A-1263 Sequence 1263, A
22	623.5	46.1	879	11	US-09-938-842A-1263 Sequence 1263, A
23	618.5	45.7	1689	16	US-10-425-114-17162 Sequence 17162, A
24	617.5	45.6	1344	16	US-10-425-114-16346 Sequence 16346, A
25	617.5	45.6	1830	16	US-10-425-115-41313 Sequence 41313, A
26	612.5	45.3	1463	16	US-10-425-114-18406 Sequence 18406, A
27	611.5	45.2	1049	16	US-10-767-795-6515 Sequence 6515, A
28	609.5	45.0	1965	16	US-10-424-599-97778 Sequence 97778, A
29	608.5	45.0	1559	16	US-10-424-599-97776 Sequence 97776, A
30	591	43.7	1404	16	US-10-424-599-133529 Sequence 133529, A
31	579.5	42.8	3076	17	US-10-437-963-48114 Sequence 48114, A
32	559.5	41.4	903	17	US-10-437-963-23863 Sequence 23863, A
33	554.5	41.0	770	17	US-10-437-963-51509 Sequence 51509, A
34	535	39.5	1103	16	US-10-424-599-117798 Sequence 117798, A
35	530.5	39.2	1276	18	US-10-425-115-128054 Sequence 128054, A
36	514	38.0	1047	16	US-10-424-599-117795 Sequence 117795, A
37	491.5	36.3	1244	18	US-10-425-115-66820 Sequence 66820, A
38	491.5	36.3	1247	16	US-10-425-114-27509 Sequence 27509, A
39	472.5	34.9	1554	16	US-10-425-114-24453 Sequence 24453, A
40	466	33.4	1195	17	US-10-437-963-5147 Sequence 5147, A
41	452.5	33.4	1183	18	US-10-739-930-2692 Sequence 2692, A
42	447	33.0	974	15	US-10-310-154-10 Sequence 10, Appl
43	441.5	32.6	1202	16	US-10-425-114-10446 Sequence 10446, A
44	429	31.7	1163	15	US-10-310-154-11 Sequence 11, Appl
45	428.5	31.7	1013	18	US-10-739-930-5185 Sequence 5185, A

#### ALIGNMENTS

RESULT 1  
US-10-424-599-85620  
; Sequence 85620, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongmei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424, 599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 85620  
; LENGTH: 1476  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_48329C.1



```
US-10-424-599-77
; Sequence 77, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5323)B
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 77
; LENGTH: 1849
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)-(1849)
; OTHER INFORMATION: unsure at all n locations
; OTHER INFORMATION: Clone ID: PAT_MRT3847_10006C.1
US-10-424-599-77

Alignment Scores:
Pred. No.:      8,93e-65      Length:      1849
Score:          654.50        Matches:      123
Percent Similarity: 73.26%    Conservative: 66
Best Local Similarity: 47.67%  Mismatches: 64
Query Match:    48.37%       Indels:       5
DB:             16           Gaps:         4

US-09-743-885a-1 (1-265) x US-10-424-599-77 (1-1849)
QY      4 LeuArgAlaAspAspSerPheIlePheGlyLeuLeuGlyValSerIleValSerPheMet 23
DB      144 ATGCATCGGAGCTTGGGCTTTTGTCTTGGGCTTATGGCAACATCATCTCTTTGGA 203
QY      24 ValPheLeuAlaProValProThrPheTYrLYsIleTYrLYsArgLYsSerSerGly 43
DB      204 GGTTCCTGCTGCTCATTTGCCAACCTTTTACCAATCTACAGAAAGAAATCTCACTGAAGT 263
QY      44 TYrGlnAlaIleProTYrMetValAlaLeuPheSerAlaGlyLeuLeuLeuTYrTYrAla 63
DB      264 TTCCACGCTCCTCTTATGTGTGTGCACTGTTCAAGTCATGCTTGGATTCTATGCT 323
QY      64 TYrLeu---ArgLYsAsnAlaTYrLeuIleValSerIleAsnGlyPheGlyCysAlaIle 82
DB      324 TTCGTGAAAAGGAAACTGCCCTCTCTTCTATTAACATTGACATTTGGAAATTTGTC 383
QY      83 GlnLeuThrTYrIleSerLeuPheLeuPheTYrAlaProArgLYsSerLYsIlePheThr 102
DB      384 GAGTCATATTACTCTTCTATCTTCTTAATTAATGAGCCCAAGGAAAGCAAGGCTTACAC 443
QY      103 ---GlyTYrLeuMetLeuLeuGlnLeuGlyAlaLeuGlyMetValMetProIleThrTYr 121
DB      444 ATCAACCTTCTCTCTTCACTGAAATGTGTGATTCGGGGCCATGCTTCTATCACTCTC 503
QY      122 LeuLeuAlaGlnGlySerHisArgValMetIleValGlyTYrIleCysAlaAlaIleAsn 141
DB      504 TACCTTCAAAGGAGCAAGAGCTTTCGTCATATGAGATGATTTGGCTTGTTCAC 563
QY      142 ValAlaValPheAlaAlaProLeuSerIleMetArgGlnValIleTYrTYrLYsSerVal 161
DB      564 ATCAGAGTCTTCTGCTCACTCTCTTCAATTAATGAGGAGTCATTAAGAAGAGAGAGCTC 623
QY      162 GlnPheMetProPheThrLeuSerLeuPheLeuThrLeuCysAlaIleMetTYrPhePhe 181
DB      624 GAATACATGCAATTTCTTTGTCATGTTTAAACATCACTGTTATAGTGGTCTTC 683
QY      182 TYrGlyPhePheLYsAspPheTYrIleAlaPheProAsnIleLeuGlyPheLeuPhe 201
```

```
DB      684 TATGGCTTCTCTCCAGGATTTATTAAGTCTCTCCAAACCGCTGGATTGTCTTC 743
QY      202 GlyIleValGlnMetLeuLeuTYrPheValTYrLYsAspSerLYsArgLYsIleAspArg 221
DB      744 GGCATTAATTCAGATGGTGTGTTGATGATGACAGAAAGCCCACTGACATAGAG 803
QY      222 LYsSerAspProValArgGlnAlaThrLYsSerLYsGlyValGlnIleIle---Ile 240
DB      804 -----GACCACTGAAGCTCAAGATTGAATGGGCATCATCTTACCTTGTGAAGAT 857
QY      241 AsnIleGluAspAspAsnSerAspAsnAlaLeuGlnSerMetGluLYsAspPhe 258
DB      858 GGAAACATGCAACCGAATCACGGTGTGCTGTGCAAAAGCTCATATGATTTTC 911

RESULT 4
US-10-425-114-32824
; Sequence 32824, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack B
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5313)B
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 32824
; LENGTH: 1102
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMFLMO17025G08_FLI
US-10-425-114-32824

Alignment Scores:
Pred. No.:      4.62e-65      Length:      1102
Score:          654.00        Matches:      121
Percent Similarity: 72.93%    Conservative: 46
Best Local Similarity: 52.84%  Mismatches: 50
Query Match:    48.34%       Indels:       12
DB:             16           Gaps:         2

US-09-743-885a-1 (1-265) x US-10-425-114-32824 (1-1102)
QY      10 SerPheIlePheGlyLeuLeuGlyValSerIleValSerPheMetValPheLeuAlaProVal 29
DB      19 GCGTTCGCTTGCTGCTTACGCAACGTCATCTCTTATATACCTTCTGAGCCCGATA 78
QY      30 ProThrPheTYrLYsIleTYrLYsArgLYsSerSerGlyLYrGlnAlaIleProTYr 49
DB      79 CGAGCTTTACCGCATCTCAAGAGCAAGTGCAGCAAGAGTTTCCAGTCGGTCTCTAC 138
QY      50 MetValAlaLeuPheSerAlaGlyLeuLeuLeuTYrTYrAlaTYrLeuArgLYsAsnAla 69
DB      139 GGGTTCGCGCTGTCAGCGCCATGCTGTGATCTTCTTACGCACTGATCAAGTCAATGAG 198
QY      70 TYrLeuIleValSerIleAsnGlyPheGlyCysAlaIleGlnLeuThrTYrIleSerLeu 89
DB      199 ACCCTTCCTATACCATCAACGCGCGGCTGCTTATGACACCATCTTACATCGTCATG 258
QY      90 PheLeuPheTYrAlaProArgLYsSerLYsIlePheThrGlyTYrLeuMet---LeuLeu 108
DB      259 TACTTCGTCTACGCGCCCAAGAAAGCAAGCGTTCACGCGCAAGATCATGCGCTCTC 318
QY      109 GlnLeuGlyAlaLeuGlyMetValMetProIleThrTYrLeuLeuAlaGlnGlySerHis 128
DB      319 AATGGCGGCTCTTTGGGATCATCTCGCTCACCCCTTCTCTTCAAGGATGATAG 378
```



```

DB      723 GCGTTGCTTCGCTCTACTAGGTAAGCATCATCTTCACTTCTCGGCCCGATA 782
QY      30 ProthPheTyrLysIleTyrLysArgLysSerSerGluGlyTyrGlnAlaIleProTyr 49
DB      783 CCGACGTTCTACCGCATCTCAAGACGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 842
QY      50 MetValAlaLeuPheSerIleGlyLeuLeuLeuTyrTyrAlaTyrLeuArgLysAla 69
DB      843 GTCGTTGCTTCGCTCTACGCGCATGCTGTGATCTTCTACGCACTGATCAAGTCCAAAG 902
QY      70 TyrLeuIleValSerIleAsnGlyPheGlyCysAlaIleGluLeuThrTyrIleSerLeu 89
DB      903 ACCCTTCCTCATCACTCAACGCGCGCGCTGCTGATGAGACATCTACGTCGATC 962
QY      90 PheLeuPheTyrLysAlaProArgLysSerLysIlePheThrGlyTyrLeuMet---LeuLeu 108
DB      963 TACTTCGCTACCGCGCCCAAGAAAGCCAGACGCTTCTACGCGCAAGATCATGATCTCTC 1022
QY      109 GluLeuGlyAlaLeuGlyMetValMetProIleThrTyrLeuLeuAlaGluLysSerHis 128
DB      1023 AATGGCGCGCTCTTGGGGTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1082
QY      129 ArgValMetIleValGlyTyrIleCysAlaAlaIleAsnValAlaValPheAlaIlePro 148
DB      1083 CCGGTTGCTGCTGCTGCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1142
QY      149 LeuSerIleMetArgGlnValIleLysThrLysSerValGluPheMetProPheThrLeu 168
DB      1143 CTCAGCATCATGAGAGAGATGATCCAGACGAGAGGCTGAGATGATGCTCTCTCTCTC 1202
QY      169 SerLeuPheLeuThrLeuCysAlaIleThrMetTyrPheTyrGlyPhePheLysValAsp 188
DB      1203 TCCCTTCCTGCTCACTCTCAAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1262
QY      189 PheTyrIleAlaPheProAsnIleLeuGlyPheLeuPheGlyIleValGlnMetLeuLeu 208
DB      1263 AATATGCTGCGCTTCCAAACATCTTGGGTTCACTTGGGCTGCGGCTGCGGATGATGCTC 1322
QY      209 TyrPheValTyrLysAspSerLysArgIleAspAspGluLysSerAspProValArgIle 228
DB      1323 TACGTTGCTACATGAGAC---AAGACGCGCGGCTGCGCGCG 1358
QY      229 AlaThrLysSerLysGluGly-----ValGluIleIleIle 240
DB      1359 ACTGCCGAGGCGAAGATGCCGGAAGCTTCTCTCACTGACGACAGACGATCTCTGCTC 1418
QY      241 AsnIle 242
DB      1419 AACATC 1424

RESULT 7
US-10-425-114-3587
; Sequence 3587, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jindong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 3587
; LENGTH: 1306
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:

```

```

; OTHER INFORMATION: Clone ID: 700332286_FLI
US-10-425-114-3587

Alignment Scores:
Pred. No.: 7,81e-65
Score: 653.00
Percent Similarity: 70.54%
Best Local Similarity: 51.04%
Query Match: 48,264
DB: 16
Gaps: 3

US-09-743-885A-1 (1-265) x US-10-425-114-3587 (1-1306)
QY      10 SerPheIlePheGlyLeuLeuGlyAsnIleValSerPheMetValPheLeuAlaProVal 29
DB      205 GCGTTGCTTCGCTCTCTTACGACGATCATCTCTTATGACCTTCTGCGCCGATA 264
QY      30 ProthPheTyrLysIleTyrLysArgLysSerSerGluGlyTyrGlnAlaIleProTyr 49
DB      265 CCGAGCTTCTACCGCATCTCAAGACGACGATGAGGAGGAGGAGGAGGAGGAGGAGGAG 324
QY      50 MetValAlaLeuPheSerIleGlyLeuLeuLeuTyrTyrAlaTyrLeuArgLysAla 69
DB      325 GTCGTTGCTGCTTACGCGCATGCTGATCTTCTACGCACTGATCAAGTCCAAAG 384
QY      70 TyrLeuIleValSerIleAsnGlyPheGlyCysAlaIleGluLeuThrTyrIleSerLeu 89
DB      385 ACCCTTCCTCATCACTCAACGCGCGCGCTGCTGATGAGACATCTTACATCTGATC 444
QY      90 PheLeuPheTyrLysAlaProArgLysSerLysIlePheThrGlyTyrLeuMet---LeuLeu 108
DB      445 TACTTCGCTACCGCGCCCAAGAAAGCCAGACGCTTCTACGCGCAAGATCATGCGCTCTC 504
QY      109 GluLeuGlyAlaLeuGlyMetValMetProIleThrTyrLeuLeuAlaGluLysSerHis 128
DB      505 AATGGCGCGCTCTTGGGGTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 564
QY      129 ArgValMetIleValGlyTyrIleCysAlaAlaIleAsnValAlaValPheAlaIlePro 148
DB      565 CCGGTTGCTGCTGCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 624
QY      149 LeuSerIleMetArgGlnValIleLysThrLysSerValGluPheMetProPheThrLeu 168
DB      625 CTAGCATCATGAGGCGGATGATCCAGACGAGAGGATGATGATGATGATGATGATGATGAT 684
QY      169 SerLeuPheLeuThrLeuCysAlaIleThrMetTyrPhePheTyrGlyPhePheLysValAsp 188
DB      685 TCCCTTCCTGCTCACTCTCAAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 744
QY      189 PheTyrIleAlaPheProAsnIleLeuGlyPheLeuPheGlyIleValGlnMetLeuLeu 208
DB      745 AATATGCTGCGCTTCCAAACGCTTGGGTTCACTTGGGCTGCGGATGATGATGATGATGAT 804
QY      209 TyrPheValTyrLysAspSerLysArgIleAspAspGluLysSerAspProValArgIle 228
DB      805 TACGTTGCTACATGAGAC-----AAGACGCGCGGCTGCGCGCT 840
QY      229 AlaThrLysSerLysGluGly-----ValGluIleIleIle 241
DB      841 GCCGTCGACGAGATGCCGGAAGCTTCTCTCACTGACGACGACGACGACGATCTCTGCTC 900
QY      242 Ile 242
DB      901 ATT 903

RESULT 8
US-10-425-115-29623
; Sequence 29623, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua

```

APPLICANT: Cao, Yongwei  
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
FILE REFERENCE: 38-21(53222)B  
CURRENT APPLICATION NUMBER: US/10/425,115  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 369326  
SEQ ID NO 29623  
LENGTH: 1415  
TYPE: DNA  
ORGANISM: Zea mays  
FEATURE:  
OTHER INFORMATION: Clone ID: MRT4577\_127027C.1  
US-10-425-115-29623

Alignment Scores:  
Pred. No.: 8,82e-65 Length: 1415  
Score: 653.00 Matches: 123  
Percent Similarity: 70.54% Conservative: 47  
Best Local Similarity: 51.04% Mismatches: 55  
Query Match: 48.26% Indels: 16  
DB: 18 Gaps: 3

US-09-743-885a-1 (1-265) x US-10-425-115-29623 (1-1415)

QY 10 SerPheIlePheGlyLeuLeuGlyAsnIleValSerPheMetValPheLeuAlaProVal 29  
Db 262 GCGTTCCGCTTCGCGCTCTGAGCAAGTCATCTCTTCATGACCTTCGCGCCGATA 321  
QY 30 ProThrPheTyrIleTyrIleValArgIleSerSerGluGlyTyrGlnAlaIleProTyr 49  
Db 322 CCGAGCTTCAAGCCGATCTACAGAGCAAGTCGAGGAGGTTTCCAGTCCGTTCCCTAC 381  
QY 50 MetValAlaLeuPheSerAlaGlyLeuLeuTyrTyrAlaTyrLeuArgIleAsnAla 69  
Db 382 GTGGTTCGCTGTTCAAGCCGATCTGATCTTCTTACGACATCACTCAACTCAAGAG 441  
QY 70 TyrLeuIleValSerIleAsnGlyPheGlyCysAlaIleGluLeuThrTyrIleSerIleu 89  
Db 442 ACCTTCTCATCAACATCAAGCCGCGCTGCTGATCTGAGACATCTACATCATGTCATG 501  
QY 90 PheLeuPheTyrAlaProArgIleSerIlePheThrGlyTyrPleuMet---LeuLeu 108  
Db 502 TACTTCTTCAAGCCGATCTACAGAGCAAGTCGAGGAGGTTTCCAGTCCGTTCCCTAC 561  
QY 109 GluLeuGlyAlaLeuGlyMetValMetProIleThrTyrLeuLeuAlaGluGlySerHis 128  
Db 562 AATGGCGCGCTTCTGGGTCATCTCTGCTCACTTCTCTTCAAGGTAAGTAAAG 621  
QY 129 ArgValMetIleValIleGlyTyrIleCysAlaIleAsnValAlaValPheAlaAlaPro 148  
Db 622 CGCGTTGCTGCTGCTGCTGATCTGCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTG 681  
QY 149 LeuSerIleMetArgGlnValIleLysThrIleSerValGluPheMetProPheThrIleu 168  
Db 682 CTAAGCATCAAGAGCGGTCATCTCAAGAGCGGTCAGTCAATGCTTCTTCTCTC 741  
QY 169 SerLeuPheLeuThrIleuCysAlaIleThrMetTrpPheTyrGlyPhePheLysAsp 188  
Db 742 TCCCTTCGCTCAACCTTCAAGCGCGCTGCTGCTGCTTCTCTTCAAGGTCATCAAGGAC 801  
QY 189 PheTyrIleAlaPheProAsnIleLeuGlyPheLeuPheGlyIleValGlnMetLeuLeu 208  
Db 802 AATATACGTCGAGCAACCAAGCTCTGCTTCACTTCTGCGGTCAGATGATGTCGTC 861  
QY 209 TyrPheValTyrIleAspSerIleArgIleAspArgIleLysSerAspProValArgIle 228  
Db 862 TACGTTCTTCAATGAC---AAGACCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 897  
QY 229 AlaThrIleSerIleGluGly---ValGluIleIleLeu 241  
Db 898 GCGGTGGCAAGATGCGCGCAAGCTTCTCTCACTGACGACGACGACGATCTCTGTCAC 957

QY 242 Ile 242  
Db 958 ATT 960

## RESULT 9

US-10-425-115-29613  
Sequence 29613: Application US/10425115  
Publication No. US20040214272A1

## GENERAL INFORMATION:

APPLICANT: La Rosa, Thomas J.  
APPLICANT: Kovalic, David K.  
APPLICANT: Zhou, Yihua

## APPLICANT: Cao, Yongwei

TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with

FILE REFERENCE: 38-21(53222)B

CURRENT APPLICATION NUMBER: US/10/425,115

CURRENT FILING DATE: 2003-04-28

NUMBER OF SEQ ID NOS: 369326

SEQ ID NO 29613

LENGTH: 1703

TYPE: DNA

ORGANISM: Zea mays

FEATURE:

OTHER INFORMATION: Clone ID: MRT4577\_127018C.1

US-10-425-115-29613

US-09-743-885a-1 (1-265) x US-10-425-115-29613 (1-1703)

QY 10 SerPheIlePheGlyLeuLeuGlyAsnIleValSerPheMetValPheLeuAlaProVal 29  
Db 525 GCGTTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 584  
QY 30 ProThrPheTyrIleTyrIleValArgIleSerSerGluGlyTyrGlnAlaIleProTyr 49  
Db 585 CCGAGCTTCAAGCCGATCTACAGAGCAAGTCGAGGAGGTTTCCAGTCCGTTCCCTAC 644  
QY 50 MetValAlaLeuPheSerAlaGlyLeuLeuTyrTyrAlaTyrLeuArgIleAsnAla 69  
Db 645 GTGGTTCGCTGTTCAAGCCGATCTGATCTTCTTACGACATCACTCAACTCAAGAG 704  
QY 70 TyrLeuIleValSerIleAsnGlyPheGlyCysAlaIleGluLeuThrTyrIleSerIleu 89  
Db 705 ACCTTCTCATCAACATCAAGCCGCGCTGCTGATCTGAGACATCTACATCATGTCATG 764  
QY 90 PheLeuPheTyrAlaProArgIleSerIlePheThrGlyTyrPleuMet---LeuLeu 108  
Db 765 TACTTCTTCAAGCCGATCTACAGAGCAAGTCGAGGAGGTTTCCAGTCCGTTCCCTAC 824  
QY 109 GluLeuGlyAlaLeuGlyMetValMetProIleThrTyrLeuLeuAlaGluGlySerHis 128  
Db 825 AATGGCGCGCTTCTGGGTCATCTCTGCTCACTTCTCTTCAAGGTAAGTAAAG 884  
QY 129 ArgValMetIleValIleGlyTyrIleCysAlaIleAsnValAlaValPheAlaAlaPro 148  
Db 885 CGCGTTGCTGCTGCTGCTGATCTGCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTG 944  
QY 149 LeuSerIleMetArgGlnValIleLysThrIleSerValGluPheMetProPheThrIleu 168  
Db 945 CTAAGCATCAAGAGCGGTCATCTCAAGAGCGGTCAGTCAATGCTTCTTCTCTC 1004  
QY 169 SerLeuPheLeuThrIleuCysAlaIleThrMetTrpPheTyrGlyPhePheLysAsp 188  
Db 1005 TCCCTTCGCTCAACCTTCAAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1064



```
Db      380 GTGTGGGGCTGTTGAGCGCATGCTGTGATCTACTAAGCGGCTGCTCAACTCCGACGAG 439
Qy      70 TyrleuValSerIleAsnGlyPheGlyCysAlaIleGluLeuThrTyrIleSerIleu 89
Db      440 TGCTCTCCATCACCATTAACTCCGCTGCTGCTGCTCATCGACCATTAATACGCGCTC 499
Qy      90 PheLeuPheTyrAlaProArgLysSerIlePheThrGlyTyrPheMetLeuLeu--- 108
Db      500 TACCTCGCTACGCGCCCAAGAGCCCAAGATGTTCAACCGCAAGCTCTCTCTCGCTC 559
Qy      109 GluLeuGlyAlaLeuGlyMetValMetProIleThrTyrLeuLeuAlaGluGlySerHis 128
Db      560 AACGTCGGCTCTTCGCTGCTATCTCTCTCAACCTCTCTCTCTCTCTCTCTCTCTCTCT 619
Qy      129 ArgValMetIleValGlyTyrPheCysAlaIleAlaIleAsnValAlaValPheAlaPro 148
Db      620 CGCATCGCTGCTCTGTGTGGGTCTGCTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 679
Qy      149 LeuSerIleMetArgGlnValIleLysThrLysSerValGluPheMetProPheThrLeu 168
Db      680 CTTAGCATCATCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 739
Qy      169 SerLeuPheLeuThrLeuCysAlaIleMetThrPhePheTyrGlyPhePheLysLysAsp 188
Db      740 TCTCTCTCTCTCACCATCAGCGCGCTGCTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCT 799
Qy      189 PheTyrIleAlaPheProAsnIleLeuGlyPheLeuPheGlyIleValGlnMetLeuLeu 208
Db      800 AATATATGCTCTTCCCAAGCTGCTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 859
Qy      209 TyrPheValTyrLysPheSerLysArgIleAspAspGluLysSerAspProValArgLys 228
Db      860 TACCCCATGTACAGAACTCTG-----ACGCCCAAG 889
Qy      229 AlaThrLysSerLysGluGlyValGlnIleIleIleAsnIleGluAspAspAsnSer 247
Db      890 GCCGTGCTGACCAAGAGCTCGAGCGCGGCGACCGGCGACCGGCGACGACCACTCC 946
```

```
RESULT 12
US-10-739-930-5395
; Sequence 5395, Application US/10739930
; Publication No. US20040216190A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
; FILE REFERENCE: 38-21(53377)B
; CURRENT APPLICATION NUMBER: US/10/739, 930
; CURRENT FILING DATE: 2003-12-18
; NUMBER OF SEQ ID NOS: 11088
; SEQ ID NO 5395
; LENGTH: 1487
; TYPE: DNA
; ORGANISM: Triticum aestivum
; FEATURE:
; OTHER INFORMATION: Clone ID: TRIA8-23APR03-CLUSTER866_1
US-10-739-930-5395
```

```
Alignment Scores:
Pred. No.: 9.07e-64 Length: 1487
Score: 644.50 Matches: 116
Percent Similarity: 76.11% Conservative: 56
Best Local Similarity: 51.33% Mismatches: 47
Query Match: 47.63% Indels: 7
DB: 18 Gaps: 2
```

US-09-743-885a-1 (1-265) x US-10-739-930-5395 (1-1487)

```
Qy      10 SerPheIlePheGlyLeuLeuGlyAsnIleValSerPheMetValPheLeuAlaProVal 29
Db      255 GCCTTGCTGCTTGGCGCTCTAGGCAAGCTATCTCTTCAAGACTTACTGCCCCCACTG 314
```

```
Qy      30 ProThrPheTyrLysIleTyrLysArgLysSerSerGluGlyTyrGlnAlaIleProTyr 49
Db      315 CCGAGCTTACACCAATCTACCGGACAGTCCAGAGGGGGTTCAGTGGCTCCCTTAC 374
Qy      50 MetValAlaLeuPheSerIleGlyLeuLeuLeuTyrTyrAlaTyrLeuArgLysAsnAla 69
Db      375 GTGTGGCGCTCTTACCGGAGATGTTGATCTACTACCGGCTGCTCAAGTCTGACGAG 434
Qy      70 TyrleuValSerIleAsnGlyPheGlyCysAlaIleGluLeuThrTyrIleSerIleu 89
Db      435 TGCTCTCCATCACCATTAACTCCGCTGCTGCTGCTCATCGACCATTAATACGCGCTC 494
Qy      90 PheLeuPheTyrAlaProArgLysSerIlePheThrGly---TyrPheMetLeuLeu 108
Db      495 TACCTCGCTACGCGCCCAAGAGCCCAAGATGTTCAACCGCAAGCTCTCTCTCTCTCTCT 554
Qy      109 GluLeuGlyAlaLeuGlyMetValMetProIleThrTyrLeuLeuAlaGluGlySerHis 128
Db      555 AATGTGGGTGCTTCTGCTTATCTCTCTCTCAACCTCTCTCTCTCTCTCTCTCTCTCTCT 614
Qy      129 ArgValMetIleValGlyTyrPheCysAlaIleAlaIleAsnValAlaValPheAlaPro 148
Db      615 CGGCTGCTATGCTCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 674
Qy      149 LeuSerIleMetArgGlnValIleLysThrLysSerValGluPheMetProPheThrLeu 168
Db      675 CTCAGCTCATCCGCTTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 734
Qy      169 SerLeuPheLeuThrLeuCysAlaIleMetThrPhePheTyrGlyPhePheLysLysAsp 188
Db      735 TCTCTCTCTCTCACCATCAGCGCGCTGCTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCT 794
Qy      189 PheTyrIleAlaPheProAsnIleLeuGlyPheLeuPheGlyIleValGlnMetLeuLeu 208
Db      795 AATATATGCTCTTCCCAAGCTGCTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 854
Qy      209 TyrPheValTyrLysPheSer-----LysArgIleAspAspGluLys 222
Db      855 TACGCCCTCTACCGCAAGCGCATGCTTCAACCGGCTCTCCCAAGAGGAGTGATCTGAT 914
```

```
RESULT 13
US-10-425-114-14414
; Sequence 14414, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jindong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaka, Jack E.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425, 114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 14414
; LENGTH: 1220
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB189-020-A7_FLI
US-10-425-114-14414
```

```
Alignment Scores:
Pred. No.: 9.97e-64 Length: 1220
Score: 643.00 Matches: 120
Percent Similarity: 70.71% Conservative: 49
```







GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: December 24, 2004, 20:51:27 ; Search time 103 Seconds  
(without alignments)  
1828.729 Million cell updates/sec

Title: US-09-743-885a-1  
Perfect score: 1353  
Sequence: 1 MAQLRADDSIFGLGNIV.....NSDNLQSMKDFSRRTSK 265

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 824507 seqs, 353394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODE=frame+pn.model -DEV=xlp  
-Q=/cgn2\_1/USPTO.epool\_p/US09743885/Runat\_23122004\_165259\_9650/app\_query.fasta\_1.455  
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=tni -MINMATCH=0.1 -LOOPEXT=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALLIGN=20 -NOR SCORE=pcr -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NOR ME-ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09743885.QCEN\_1\_128.qrunat\_23122004\_165259\_9650 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA.\*  
1: /cgn2\_6/prodata/1/ina/5A\_COMB.seq.\*  
2: /cgn2\_6/prodata/1/ina/5B\_COMB.seq.\*  
3: /cgn2\_6/prodata/1/ina/6A\_COMB.seq.\*  
4: /cgn2\_6/prodata/1/ina/6B\_COMB.seq.\*  
5: /cgn2\_6/prodata/1/ina/PCITUS\_COMB.seq.\*  
6: /cgn2\_6/prodata/1/ina/backfill1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysts of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	206.5	15.3	272	4	US-09-313-294A-448 Sequence 448, App
2	195.5	14.4	1314	4	US-09-589-360B-54 Sequence 54, Appl
3	107.5	7.9	2160	2	US-08-902-853-4 Sequence 4, Appl
4	107	7.6	759	4	US-09-248-796A-12834 Sequence 12834, A
5	103.5	7.6	503	4	US-09-270-767-8810 Sequence 8810, Ap
6	103.5	7.6	503	4	US-09-270-767-24092 Sequence 24092, A
7	103	7.6	640681	4	US-09-790-988-1 Sequence 1, Appl
8	101.5	7.5	1188	4	US-09-134-000C-2036 Sequence 2036, Ap
9	101	7.5	1280	3	US-09-199-737-1 Sequence 1, Appl
10	101	7.5	1280	3	US-08-900-230-1 Sequence 1, Appl
11	101	7.5	1280	3	US-09-058-333A-1 Sequence 1, Appl
12	100.5	7.4	1389	4	US-09-328-352-1371 Sequence 1371, Ap

13	99	7.3	864	4	US-09-543-681A-1110 Sequence 1110, Ap
14	99	7.3	1146	3	US-09-134-001C-1135 Sequence 1135, Ap
15	98.5	7.3	1204	4	US-09-711-164-238 Sequence 238, App
16	98	7.2	885	4	US-09-328-352-3427 Sequence 3427, Ap
17	98	7.2	978	4	US-09-328-352-1634 Sequence 1634, Ap
18	97.5	7.2	1449	4	US-09-328-352-3658 Sequence 3658, Ap
19	97	7.2	699	4	US-09-543-681A-686 Sequence 686, App
20	97	7.2	1323	4	US-09-248-796A-6305 Sequence 6305, Ap
21	97	7.2	1664976	4	US-08-916-421B-1 Sequence 1, Appl
22	97	7.2	1664976	4	US-08-692-570-1 Sequence 1, Appl
23	96.5	7.1	13856	4	US-08-956-171B-31 Sequence 31, Appl
24	96.5	7.1	13856	4	US-08-781-986A-31 Sequence 31, Appl
25	96	7.1	12395	4	US-09-270-767-11401 Sequence 11401, A
26	95.5	7.1	297	4	US-09-313-294A-7017 Sequence 7017, Ap
27	95.5	7.1	3342	4	US-08-961-527-206 Sequence 206, App
28	95.5	7.1	1830121	4	US-09-557-884-1 Sequence 1, Appl
29	95.5	7.1	1830121	4	US-09-643-990A-1 Sequence 1, Appl
30	95.5	7.1	1830121	4	US-10-329-960-1 Sequence 1, Appl
31	95	7.0	1230025	4	US-09-198-452A-1 Sequence 1, Appl
32	93.5	6.9	1058	4	US-09-016-434-1206 Sequence 1206, Ap
33	93.5	6.9	1296	4	US-09-489-039A-1107 Sequence 1107, Ap
34	93.5	6.9	1536	4	US-09-252-991A-11740 Sequence 11740, A
35	93.5	6.9	1548	4	US-09-489-039A-1037 Sequence 1037, Ap
36	93.5	6.9	1537	4	US-09-252-991A-15015 Sequence 15015, A
37	92.5	6.8	1910	4	US-09-944-807-1 Sequence 1, Appl
38	92.5	6.8	2163	4	US-09-134-000C-2226 Sequence 2226, Ap
39	92.5	6.8	2631	4	US-09-023-655-1083 Sequence 1083, Ap
40	92.5	6.8	1664976	4	US-08-916-421B-1 Sequence 1, Appl
41	92.5	6.8	1664976	4	US-09-692-570-1 Sequence 1, Appl
42	92	6.8	1977	4	US-09-614-891-5 Sequence 5, Appl
43	91.5	6.8	1014	4	US-10-061-943A-3 Sequence 3, Appl
44	91.5	6.8	1014	4	US-09-016-434-750 Sequence 750, App
45	91.5	6.8	1301	2	US-08-467-948A-7 Sequence 7, Appl

## ALIGNMENTS

RESULT 1  
US-09-313-294A-448  
; Sequence 448, Application US/09313294A  
; Patent No. 6476212  
; GENERAL INFORMATION:  
; APPLICANT: Ialsudi, Raghunath V.  
; APPLICANT: Ito, Laura Y.  
; APPLICANT: Sherman, Bradley K.  
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR  
; FILE REFERENCE: PL-0017 US  
; CURRENT APPLICATION NUMBER: US/09/313, 294A  
; CURRENT FILING DATE: 1999-05-14  
; NUMBER OF SEQ ID NOS: 7600  
; SOFTWARE: PERL Program  
; SEQ ID NO 448  
; LENGTH: 272  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; NAME/KEY: misc\_feature  
; OTHER INFORMATION: Incyte ID No. 6476212 700549251H1  
; US-09-313-294A-448

## Alignment Scores:

Pred. No.: 4.5e-17 Length: 272  
Score: 206.50 Matches: 44  
Percent Similarity: 73.26% Conservative: 19  
Best Local Similarity: 51.16% Mismatches: 22  
Query Match: 15.26% Indels: 2  
DB: 4 Gaps: 1

US-09-743-885A-1 (1-265) x US-09-313-294A-448 (1-272)

QY 80 CysAlIleGluLeuThrTyrIleSerLeuPheLeuPheTyrIlaProArgIysSerIys 99  
Db 15 TGGCATCGAGACCATCTACATCGTCATGTACTTGGTCACGCGCCCAAGAACCAAG 74

[illegible]

```

TELEX:
/ INFORMATION FOR SEQ ID NO: 4:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 2160 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ IMMEDIATE SOURCE:
/ LIBRARY: HNT2NOT01
/ CLONE: 493014
/ US-08-902-853-4

Alignment Scores:
Pred. No.: 0.00279 Length: 2160
Score: 107.50 Matches: 71
Percent Similarity: 38.78% Conservative: 31
Best Local Similarity: 27.00% Mismatches: 97
Query Match: 7.95% Indels: 65
DB: 2 Gaps: 14

US-09-743-885a-1 (1-265) x US-08-902-853-4 (1-2160)
Qy 1 MetAlaGluLeuArgLAlaAspLeuSerPheIlePheGlyLeuLeuGlyAlaIleVal 20
Db 494 CTCAGAAGATTCGAGAGAGCCAGCTGAGATTCACATTTTACTGATTCGCTTTCATT 550
Qy 21 SerPheMetValPheLeuAlaProValProThrPheTyrIleTyrIleTyrIleSer 40
Db 551 GCCGCGATGCGCGCTGATGATTAACCTGCTTCTATGACATG-----AAGAAAGTT 604
Qy 41 SerGluGlyTyr-----GlnAlaIlePro-----TyrMetValAla 52
Db 605 TGGGAGGATATCCCATACAGACACTATCCCTCCAGATTTGATCACTCAATGATGA 664
Qy 53 LeuPhe-SerAlaGlyLeuLeuLeuTyrTyrAlaTyrLeuArgLysAlaIleTyrLeuI 72
Db 665 CTTTCCTCTACTGCT----- 680
Qy 72 eValSerIleAsnGlyPheGlyCysAlaIleGluLeuThrTyrIleSer----- 88
Db 681 -----CCCTGCTCTTCAGCATTCCTGATGCTCAGCGAAAGAT 721
Qy 89 -----LeuPheLeuPheTyrAlaProArgLysSerIlePheThrGlyTyr 104
Db 722 TTCAGAAGACAGATTCACCATGTCGACCATTCATTCATTCAGCTTTCTGCTTT 781
Qy 104 PheMetLeuLeuGlyLeuGlyAla-----LeuGlyMetValMetProIleThrTyr 121
Db 782 GCCAATTACATCGGAGCTGGACTTAATCATGCTCTGATCATCTTCGATTCCTG 841
Qy 121 rLeuLeuAlaGlyLysSerHisArgValMetIleValGlyTyrIleCysAlaAlaIleAs 141
Db 842 -CTG-----GAGTCAGCCAAAGATGTTTAACAGCGGAGTGAAGAACCTCGAACAA 894
Qy 141 nVal-----AlaValPheAlaAlaProLeuSerIleMetArgGlnValIleLysThrLys 160
Db 895 CACTTCATCATGCTTGGCTGCTTTTATTCATCACCCGACGCTGATC----- 943
Qy 160 rValGluPheMetProPheThrLeuSerLeuPheLeuThrLeuCysAlaIleThrMetTrp 180
Db 944 -----CTGCCCTTCTGATC---CTGCATTCGACCTGCTGTATGACCATCGAGCT 990
Qy 180 ePheTyrGlyPhePheLeuLysAspPheTyrIleAlaPheProAsnIleLeuGlyPheLe 200
Db 991 CTATCTGCTCTTCTGCTATTCATTCCTCAATTCATCATGAGGAGTTCACAGCTGCT 1050
Qy 200 u-----PheGlyIleValGluMetLeuLeuTyrPheValTyrLysAspSe 215
Db 1051 GCATATCTTCTGGGCTTACTCTCATTTTGGCATGGCCACAAAGTTCTTA-----ACTGG 1104
Qy 215 rLysArgLysAspAspGluLysSerAspProValArgGluAlaThrLysSerLysGlu 235
Db 1105 AAGGCTGTGAAGATGAAGACGAGTGAC-----CGGAGAAAGAAAGAGAGCTCAGAGGG 1158

```

```

Qy 235 yValGlu 237
Db 1159 GAGAGAG 1165

RESULT 4
US-09-248-796A-12834
/ Sequence 12834, Application US/09248796A
/ Patent No. 6747137
/ GENERAL INFORMATION:
/ APPLICANT: Keith Weinstein et al
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
/ FILE REFERENCE: 107196, 132
/ CURRENT APPLICATION NUMBER: US/09/248, 796A
/ PRIOR FILING DATE: 1999-02-12
/ PRIOR APPLICATION NUMBER: US 60/074, 725
/ PRIOR FILING DATE: 1998-02-13
/ PRIOR APPLICATION NUMBER: US 60/096, 409
/ NUMBER OF SEQ ID NOS: 28208
/ SEQ ID NO 12834
/ LENGTH: 759
/ TYPE: DNA
/ ORGANISM: Candida albicans
US-09-248-796A-12834

Alignment Scores:
Pred. No.: 0.00647 Length: 759
Score: 107.00 Matches: 58
Percent Similarity: 42.92% Conservative: 42
Best Local Similarity: 24.89% Mismatches: 92
Query Match: 7.91% Indels: 42
DB: 4 Gaps: 9

US-09-743-885a-1 (1-265) x US-09-248-796A-12834 (1-759)
Qy 43 GlyTyrGlnAlaIleProTyrMetValAlaLeu-----PheSerAlaGlyLeuLeu----- 59
Db 49 GGGCTCAATCGGATATTTGCAATGTCCTTGGCTTTCAGCTGGGTTTGGCTGAT 108
Qy 60 LeuTyrTyrAlaTyrLeuArgLysAsnAlaTyrLeuIleValSerIleAsnGlyPheGly 79
Db 109 GTAGATATTAATTAACAGCGGTGACCTTTGGCTTGATGTTCCATA----- 156
Qy 80 CysAlaIleGluLeuThrTyrIleSerLeuPheLeuPheTyrAlaProArgLysSerLys 99
Db 157 -----CTCAATCTTATCTATTTAGTTACTACTCATGCTCACTATTCACAA 210
Qy 100 IlePheThrGlyTyrLeuMetLeuLeuGlyAlaLeuGlyMetValMetProIle 119
Db 211 AACTTTTCACTTCCTGGTAAATCATTTTA-----GTGCCGAGTTCAATTTCTTTG 261
Qy 120 ThrTyrLeuLeuAlaGlyLysSerHisArgValMetIleValGlyTyrIleCysAlaAla 139
Db 262 TTTTATCTTTCAGCAGTGAAGCAATTCAGACCGTATTCCTTCGGAAGTTGTGA--- 318
Qy 140 IleAsnValAlaValPheAlaAlaProLeuSerIleMetArgGlnValIleLysThrLys 159
Db 319 ---GATTATGGCAGTTACTCATCATCGCATCAGCATCTCAAGGCTTTAATA----- 366
Qy 160 SerValGluPheMetProPheThrLeuSerLeuPheLeuThrLeuCysAlaIleThrMetTrp 179
Db 367 -----CCATTTACTTATTTCAACTGAGTTATTTGTTCTCTACTGTTTGA 411
Qy 180 PhePheTyrGlyPhe-----PheLys-LysAspPhe 189
Db 412 TTATTTTGGGTTATAGCTTATTCACAAAGTCAGTAGTCGGAGTTCAAGTCAATCTTT 471
Qy 189 eTyrIleAlaPheProAsnIleLeuGlyPheLeuPheGlyIleValGluMetLeuLeuTyr 209
Db 472 TTACTCGACAGATTCAGGTTGTGTCTATTTTACG-----ATTTCGCTTAA 519

```

```

Oy 209 rPheValTyrLysAspSerLysArgLlePheAspGluLysSerAspProValArgGlu 229
      ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 520 CCATTGGAAAAAATATCCGGTA-ACTGATCCAGTAAGTGAAGCTGCTTAAGTCAATGC 578
Oy 229 aThrLysSerLysGluGlyValGlu-----IleIleIleAsnLleGluAsp 245
      ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 579 TGAAGTTCATCGTAGCGTATGAGAAATGATGATGCTCTTAAGATGTCATCTGTGTGATAA 638
Oy 245 pAsnSerAspAsnAlaLeuGlnSerMetGluLysAsp 257
      ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 639 TGAGCAACTGTTGGTTGGCTTCATCCAGAGAAGAT 675

RESULT 5
US-09-270-767-8810
; Sequence 8810, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8810
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-8810

Alignment Scores:
Pred. No.: 0.00945 Length: 503
Score: 103.50 Matches: 26
Percent Similarity: 47.87% Conservative: 19
Best Local Similarity: 27.66% Mismatches: 40
Query Match: 7.65% Indels: 9
DB: 4 Gaps: 1

US-09-743-885A-1 (1-265) x US-09-270-767-8810 (1-503)
Oy 133 ValGIYTPILeCyAlaAlaIleAenValAlaValPheAlaLAProLeuSerIleMet 152
      ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 135 CTGGGAATGTTGATCATCCGGCATCTCGTTGGATGCTGGCGTCTCCCTGTCGATCTG 194
Oy 153 ArgGlnValIleLysThrLysSerValGluPheMetProPheThrLysSerLeuPheLeu 172
      ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 195 CCGAAATCATCATGAGAGAAGAGACCGAGGAGATCCGTTCCCGATTATCTTGGCCGT 254
Oy 173 ThrLeuCyAlaIleThreTrpPhePheTyrGlyPhePheLysAspAspPheTyrIleAla 192
      |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 225 AATCTGGTGCGCATTTTCTCGACGCTGTAATGCCATCTCCATCAAGAAATACTGTGATGCTG 314
Oy 193 PheProAsnIleLeuGlyPheLeuPheGlyIleValGlnMetLeuDeuTyrPheValTyr 212
      ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 315 CTGCAAGATCTTTGTGCTGTGTCGTCGCGGCGGCGATTCACGTCCTCCATGTTCCGATTATT 374
Oy 213 -----LysAspSerLysArg 217
      ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 375 CCGAACAAACCGGCTGCCGAGAACCCCAAGCGACAGCAAGAAG 416

RESULT 6
US-09-270-767-24092
; Sequence 24092, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 24092

```

```

; LENGTH: 503
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-24092

Alignment Scores:
Pred. No.: 0.000945 Length: 503
Score: 103.50 Matches: 26
Percent Similarity: 47.87% Conservative: 19
Best Local Similarity: 27.66% Mismatch: 40
Query Match: 7.65% Indels: 9
DB: 4 Gaps: 1

US-09-743-885A-1 (1-265) x US-09-270-767-24092 (1-503)
QY 133 VALGIYTPILIECYALALALALLENVALALAVAlPhALALALProLeuSerILleMe 152
DB ::::: ::::: ::::: ::::: ::::: :::::
135 CTGGGAATGTTGATATCACCGGCATCTGTTTGAAATGGGCGCTCCCTGTCGATCTG 194
QY 153 ATGTAVALIIElyTThriLySerValGIuPhMeTProPhEThrPhEThrLeuPhLeu 172
DB ::::: ::::: ::::: ::::: ::::: :::::
195 CCGAATAATCATCGAAGAAGACACCGAGAAATGCCCTTCCCGATTATCTTTGCCGGR 254
QY 173 ThriLeuCYaIaThMeTTPhePhETyrgLyIphepHeLyALySaPhETyrlIeAla 192
DB 255 AATCTGGAGCGCATTTTCTCGAGCGCTGATGCCATCTCCATCAAGAATCTGTGAAGTG 314
QY 193 PheProAenILleLeuGIyPheLeuPhEGyILeValGIuMeTLeuLeuTyrrPheValTYr 212
DB 315 CTGCGAATCTTTTCTGCTGCTGCTGCTGCTGCGCGCGCATTCACCTCTCCATGTTGCTATTAT 374
QY 213 -----LySaPSeTlySaY 217
DB 375 CCCAACAAACGCGTGCAGAACCCCAAGACAGACAGAAG 416

RESULT 7
US-09-790-988-1
; Sequence 1, Application US/09790988
; Patent No. 6632935
; GENERAL INFORMATION:
; APPLICANT: SHIGENOH, SHUJI
; APPLICANT: MATANABE, HIDEMI
; APPLICANT: HATTORI, MASAHIRA
; APPLICANT: SAKAKI, YOSHITUKI
; TITLE OF INVENTION: GENOME DNA OF BACTERIAL SYMBIONT OF APHIDS
; FILE REFERENCE: 081356/0159
; CURRENT APPLICATION NUMBER: US/09/790,988
; CURRENT FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: JP2000-107160
; PRIOR FILING DATE: 2000-04-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 640681
; TYPE: DNA
; ORGANISM: Buchnera sp.
US-09-790-988-1

Alignment Scores:
Pred. No.: 63.4 Length: 640681
Score: 103.00 Matches: 55
Percent Similarity: 40.64% Conservative: 47
Best Local Similarity: 21.91% Mismatch: 90
Query Match: 7.61% Indels: 59
DB: 4 Gaps: 14

US-09-743-885A-1 (1-265) x US-09-790-988-1 (1-640681)
QY 3 GINLEUAAGAlaAspAlaSerHeILePhEGyLeuLeuClYAsnIleValSerPhe 22
DB ::::: ::::: ::::: ::::: ::::: :::::
176185 AGATTATCGCGATCCACTTAT-----AAATTATCAATTAT 1762222222
QY 23 MetValPheLeuAlaProValProThrPheTyrlYsIleTyrlYsArGIySersSergIu 42

```



Pred. No.:	0.00915	length:	1280
Score:	101.00	Matches:	46
Percent Similarity:	40.38%	Conservative:	40
Best Local Similarity:	21.60%	Mismatches:	67
Query Match:	7.46%	Indels:	60
DB:	3	Gaps:	10

[illegible]

```

1      SOFTWARE: Patent In Release #1.0, Version #1.30
2      CURRENT APPLICATION DATA:
3      APPLICATION NUMBER: US/08/900,230
4      FILING DATE: 23-JUL-1997
5      CLASSIFICATION: 435
6      ATTORNEY/AGENT INFORMATION:
7      NAME: White, John P.
8      REGISTRATION NUMBER: 28,678
9      REFERENCE/DOCKET NUMBER: 52241-C/JPW/ADM
10     TELECOMMUNICATION INFORMATION:
11     TELEPHONE: 212-278-0400
12     TELEFAX: 212-391-0525
13     INFORMATION FOR SEQ ID NO: 1:
14     SEQUENCE CHARACTERISTICS:
15     LENGTH: 1280 base pairs
16     TYPE: nucleic acid
17     STRANDEDNESS: single
18     TOPOLOGY: linear
19     MOLECULE TYPE: other nucleic acid
20     HYPOTHEetical: NO
21     ANTI-SENSE: NO
22     JS-08-900-230-1

```

US-09-743-885A-1 (1-265) x US-08-900-230-1 (1-1280)

Qy 9 LeuSerPheIlePheGlyLeuLeuGlyValnIleValSerPheMetValPheLeuAlaPro 28  
Db 138 CTTATCTTCTCTGTGGGCAATGTGGCAATAGCGGTGGTCTGTGGCTATCTGCAGCCT 137  
Qy 29 ValProThrPheTyrlleTyrlleTyrlleValArgLysSerSerGluGly----- 43  
Db 198 GGCCCAAGT-----GCTGGGAGAGAGCCAGACAGTACACAGATCTTCTATCTTC 248  
Qy 44 -----TyrlnAlaIleProTyrlMetValAlaLeu 53  
Db 249 AACTGGCCGCTGGCGACCTTTCCTTACCTCGTGTCTGCTGGCCCTTCCAGGCAAGCATC 308  
Qy 54 PheSerAlaGlyLeuLeuLeuTyrlzAlaTyrlleuArgLysAlaAlaTyrlleuIleVal 73  
Db 309 TACACACTAGATAGCTCGCTCTTGTGGGCTTTCGTGTGCAlaAGACGATACATCTGTATC 368  
Qy 74 SerIleAsnGlyPheGlyCysAlaIleGluLeuThrTyrlleSerLeuPheLeuPheTyrl 93  
Db 369 TACTTACACATGATATGCCAGACAGCTTACCCGTGGCGGCGTCTCCCTGGACAGTACTG 428  
Qy 94 Ala-----ProArgLysSerLysIlePhe 101  
Db 429 GCTGTGCGGACCCACTGCGCTCCAGAGCCCTGCGCACCCGCGCCACGCGCGCGCC 488  
Qy 102 ThrGly-----TrpLeuMetLeuLeuGluGlyAlaLeuGlyMetValMetProIle 119  
Db 489 GTGGGGCTGTGTGGTGTG-----CTGGCGGCTCTCTTTTCCGGCCCTTACTTA 538  
Qy 120 ThrTyrl-----LeuLeuAlaGluLysSerHisArgValMetIleValGlyTrpIleCys 137  
Db 537 AGCTATTACGCGACGCTGTGGCTACGCGCGCTCCAGCTCTGTGCGCGCGCTGGGAAGAC 596  
Qy 138 Ala-----AlaIleAsnValAlaValPheAlaAla-----ProLeuSer 156  
Db 597 GCCCGGCGCGCGCGCTGACGCTGACCTTCGCGCGGCGCTACCTGCTCCGCTGAGC 656  
Qy 151 IleMetArgIleValIleLysThrLysSerValGluPheMetProPheThrLeuSerLeu 176  
Db 657 GTGGTG-----AGCTAGGCTTAC 674



QY 171 PheLeuThreucysAlaThrMetTrpPhePheTyrgly 183  
 DB 675 GAGCGACGCTATGT---TTCTATGGCGCGCGTGGGT 710

RESULT 11  
 US-09-058-333A-1  
 ; Sequence 1, Application US/0905833A  
 ; Patent No. 6368812  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Bard, Jonathan A  
 ; APPLICANT: Borowsky, Bech  
 ; APPLICANT: Smith, Kelli E  
 ; TITLE OF INVENTION: DNA ENCODING GALANIN GALR3 RECEPTORS  
 ; TITLE OF INVENTION: AND USES THEREOF  
 ; NUMBER OF SEQUENCES: 65  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Cooper & Dunham LLP  
 ; STREET: 1185 Avenue of the Americas  
 ; CITY: New York  
 ; STATE: New York  
 ; COUNTRY: U.S.A.  
 ; ZIP: 10036  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patentin Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/058,333A  
 ; FILING DATE:  
 ; CLASSIFICATION: 435  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: White, John P  
 ; REGISTRATION NUMBER: 28,678  
 ; REFERENCE/DOCKET NUMBER: 52241-E/JPM/KDB  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 212 278 0400  
 ; TELEFAX: 212 391 0525  
 ; INFORMATION FOR SEQ ID NO: 1:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 1280 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: DNA  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: 63..1172  
 ; US-09-058-333A-1

Alignment Scores:  
 Pred. No.: 0.00815 Length: 1280  
 Score: 101.00 Matches: 46  
 Percent Similarity: 40.38% Conservative: 40  
 Best Local Similarity: 21.60% Mismatches: 67  
 Query Match: 7.46% Indels: 60  
 DB: 3 Gaps: 10

US-09-743-885A-1 (1-265) x US-09-058-333A-1 (1-1280)

QY 9 leuSerPheIlePheglyLeuLeuGlyAsnIleValSerPheMetValPheLeuAlaPro 28  
 DB 138 CTATCTTCTGTGGGCGATGGGCGAATGGGCTGTGGCTGTGCTACTGACGCT 197

QY 29 ValProThrPheTyrlsIleTyrlsArgLysSerSerGluGly----- 43  
 DB 198 GGGCCAAAGT-----GCCTGGCAGAGCAAGCAAGCAAGATCTTCATCTC 248

QY 44 -----TyrGlnAlaIleProTyrlMetValAlaLeu 53  
 DB 249 AACTGGCCGTGGCGGACCTTTGGCTTCATCTGTGCTGCTGCCCTTCAGGAGCCATC 308

QY 54 PheSerIleaglLeuLeuLeuTyrlTyrlAlaTyrlLeuArgLysAsnAlaTyrlLeuIleVal 73

DB 309 TACACCTGAGANGCCTGGCTCTTGTGGGGCTTGTGTGCAAGACGGACATCTGCATC 368  
 QY 74 SerIleAsnGlyPheglyCysAlaIleGluLeuThrTyrlIleSerLeuPheLeuPheTyrl 93  
 DB 369 TACCTCACCATGATATCCACAGACGCTTACACCTTGCGCGGCTTCCCTTGACAGATACCTG 428

QY 94 Ala-----ProArgLysSerLysIlePhe 101  
 DB 429 GCTGTGCGGACCACTGCGCTCCAGAGCCCTGCGGACCCCGGCAAGCGGCGCGCC 468

QY 102 ThrGly-----TrpLeuMetLeuLeuGlyAlaLeuGlyMetValMetProIle 119  
 DB 489 GTGGGGCTGTGGCTG-----CTGGCGGCTCTCTTTCCGGCGGCTTACTA 536

QY 120 ThrTyrl-----LeuLeuAlaGluGlySerHisArgValMetIleValGlyTrpIleCys 137  
 DB 537 AGCTATTACGCGACGCTGCGCTACAGCGCGGCTGCTGCTGCTGCTGCGCGGAGAG 596

QY 138 Ala-----AlaIleAsnValAlaValPheAlaIle-----ProLeuSer 150  
 DB 597 GCGCGCGCGCGCGCGCTGGAHCGTGGCACCTTGCGCGGCGGCTACCTGCTGCGGCGG 656

QY 151 IleMetArgGlnValIleLeuTyrlsSerValGluPheMetProPheThrLeuSerLeu 170  
 DB 657 GTGGTG-----AGCTGGGCTAC 674

QY 171 PheLeuThreucysAlaThrMetTrpPhePheTyrgly 183  
 DB 675 GAGCGACGCTATGT---TTCTATGGCGCGCGTGGGT 710

RESULT 12  
 US-09-328-352-1371  
 ; Sequence 1371, Application US/09328352  
 ; Patent No. 6562958  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gary L. Breton et al.  
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
 ; FILE REFERENCE: GTC99-03PA  
 ; CURRENT APPLICATION NUMBER: US/09/328,352  
 ; CURRENT FILING DATE: 1999-06-04  
 ; NUMBER OF SEQ ID NOS: 8252  
 ; SEQ ID NO 1371  
 ; LENGTH: 1389  
 ; TYPE: DNA  
 ; ORGANISM: Acinetobacter baumannii  
 ; US-09-328-352-1371

Alignment Scores:  
 Pred. No.: 0.0107 Length: 1389  
 Score: 100.50 Matches: 59  
 Percent Similarity: 39.92% Conservative: 40  
 Best Local Similarity: 23.79% Mismatches: 91  
 Query Match: 7.43% Indels: 59  
 DB: 4 Gaps: 13

US-09-743-885A-1 (1-265) x US-09-328-352-1371 (1-1389)

QY 11 PheIlePheglyLeuLeuGlyAsnIleValSerPheMetValPheLeuAlaProValPro 30  
 DB 670 TTTTTCATCGAGACCTTTTACGTATTTAGTGTTCGATTTGGTGGCTGCTATTGGAA 729

QY 31 ThrPheTyrlsIleTyrlsArgLysSerSerGluGlyTyrGlnAlaIleProTyrlMet 50  
 DB 730 ACT-----CAATCCTTAAACGCTCAAGCAAGAAACAGATCAGCGAATCAGGATG 783

QY 51 ValAlaLeuPhe-----SerAla 56  
 DB 784 TTTGCTTATTATTAAACATTATCTTAAAGAGCCTTACGTATTTATTTCTAATGCTGCT 843

QY 57 GlyLeuLeuLeuTyrlTyrlAlaTyrlLeuArgLysAsnAlaTyrlLeuIleValSerIleAsn 76



:	ORGANISM:	Escherichia coli
:	FEATURE:	
:	NAME/KEY:	CDS
:	LOCATION:	(1)...(1704)
:	US-09-711-164-238	
 Alignment Scores:		
Pred. No.:	0.026	Length: 1704
Score:	98.50	Matches: 56
Percent Similarity:	37.55%	Conservative: 48
Best Local Similarity:	20.22%	Mismatches: 86
Query Match:	7.28%	Indels: 87
DB:	4	Gaps: 14
 US-09-743-885A-1 (1-265) x US-09-711-164-238 (1-1704)		
Oy	23 MetValPheLeuAlaProValProThrPheTyrIleTyrIysArgLySerSerGlu	42
Dd	1 ATGGCTTTTGATCGAAAGAAATTGTAGAAGAGATATTACCAAAAAAATTGACTTT	60
Oy	43 GltYrGlnAlaIleProTyrMetValAlaLeuPhe-----SerAlGlyLeuLeu	59
Dd	61 ACGATCTTGGCATTTCATTAATTCCTCATTTTTTTCTTCTTGTAGAGGGCTTCTC	120
Oy	60 LeuTyrTyr-----AlaTyrLeuArgLysAsnAlaTyrLeuIleValSerIleAsn	76
Dd	121 TTTCGAATTTTTCACGCCGCTTCACAAMAGATAATTGTAT-----	162
Oy	77 GlyPheGlyCySaIaIleGluLeuThrTyrIleSerLeuPheLeuPheTyrAlaProArg	96
Dd	163 -----GATGCTCAATGMAAATTGATTACTATAAACATTTTATTT-----	204
Oy	97 LysSerIysIlePheThrGlyTyrPheLeuMetLeuGluLeuGlyAlaLeuGlyMetVal	116
Dd	205 -----TCACTGGAACCCGATGGTACTC-----ATATCGCTGTTGCAGTATTT	249
Oy	117 MetProIleThrTyrLeuLeuAlaGluGlySerHisArgVal-----	130
Dd	250 ATTGCATCTGTTTTTATA-----CAGCATGAAATTAAGCTTATTTAACAATTA	297
Oy	131 MetIleValGlyTyrPheCySaIaIle-----	140
Dd	298 TTACTCTTACATGATGATTTGTAACAATAACAGATGTGGCTTAATAACGCTTAGAC	357
Oy	141 AsnValAlaValPheAlaAlaProLeuSerIleMetArgGlnValIleLysThrLysSer	160
Dd	358 AATATGACATGAAATATATTTGTTAAATATACTATATATATCTT-----	402
Oy	161 ValGluPheMetProPheThrLeuSerLeuPheLeuThrLeuCysAlaIleMetCitrPhe	180
Dd	403 -----TTTGGGGCGATTTTATTTGCTGCTGTTTATGTTTATCAAAAATCCCTCTTTTC	456
Oy	181 PheTyrGlyPhePheIysLysAspPheTyrIleAlaPheProenIIeLeuGlyPheLeu	200
Dd	457 CACCTTAATAAGATAAACAAT-----ATTCCAATGATTTATCGGCATGT	501
Oy	201 PheGlyIleValGlnMetLeu-----LeuTyrPheVal	211
Dd	502 ATTGCATTAGTTTCCGCAATTAATAAGCCATCTACTCAACCGCTTATATATTTATGTG	561
Oy	212 TyrIysAspSerLysArgIleAspAspGluLysSerAspProValArgGluAlaThrLys	231
Dd	562 TTT-----	564
Oy	232 SerIysGluGlyValGluIleIleIleAsnIleGluAspAsnSerAspAsnAla---	250
Dd	565 GCCCGGACGGCGGTAGAGATTGACATGATATA--TCTGAAGTTTCAGACATCGCATAT	621
Oy	251 -----LeuGlnSerMetGluLysAspPheSerArgLeuArgThrSerLys	265
Dd	622 GTAGCTGTAAAGAATATAGAGATCATTTTCGTTTAAATGATGATAAAAA	672

Mon Dec 27 09:34:22 2004

us-09-743-885a-1.rni

Page 10

Search completed: December 24, 2004, 22:23:41  
Job time : 232 secs

---

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: December 24, 2004, 20:44:17 : Search time 589 Seconds  
(without alignments)  
2361.794 Million cell updates/sec

Title: US-09-743-885a-1

Perfect score: 1353  
Sequence: 1 MAQLRADLDFIFGLGNIV.....NSDNALQSMKDFSLRTSK 265

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+pin.model -DEV=xlp  
-O=/cgn2.1/USPTO\_epool\_p/US09743885/runat\_23122004\_165258\_9626/app\_query.fasta\_1.455  
-DB=N Geneseq\_23Sep04 -QFMT=faextap -SUFIX=ring -MINMATCH=0.1 -LOOPEL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09743885@CGN\_1\_1.885@runat\_23122004\_165258\_9626 -NCPU=6 -ICPU=3  
-NO\_WMAP -LANG=QUEYRY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOC  
-DEV TIMEOUT=120 -WAPN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: N Geneseq\_23Sep04:\*  
2: geneseqn19808:\*  
3: geneseqn19908:\*  
4: geneseqn20008:\*  
5: geneseqn20018:\*  
6: geneseqn20028:\*  
7: geneseqn20028:\*  
8: geneseqn20038:\*  
9: geneseqn20038:\*  
10: geneseqn20038:\*  
11: geneseqn20038:\*  
12: geneseqn20048:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1353	100.0	1205	3	AAZ35493
2	1353	100.0	1205	3	AAZ50200
3	1309	96.7	847	3	AAZ35498
4	1309	96.7	847	3	AAZ50207
5	704.5	52.1	900	3	AAZ47558
6	703.5	52.0	902	3	AAZ40321

7	623.5	46.1	879	6	ABZ13458	Abz13458 Arabidops
8	623.5	46.1	1263	3	AAZ34308	Aaz34308 Arabidops
9	622	46.0	870	12	ADN73900	Adn73900 Thale cre
10	622	46.0	1177	3	AAZ51633	Aaz51633 Arabidops
11	622	46.0	1189	3	AAZ41129	Aaz41129 Arabidops
12	621.5	45.9	1219	3	AAZ49337	Aaz49337 Arabidops
13	621.5	45.9	1221	3	AAZ39463	Aaz39463 Arabidops
14	618	45.7	985	3	AAZ46920	Aaz46920 Arabidops
15	590	43.6	1252	3	AAZ47147	Aaz47147 Arabidops
16	564.5	41.7	1159	3	AAZ33952	Aaz33952 Arabidops
17	564.5	41.7	1159	3	AAZ47391	Aaz47391 Arabidops
18	548.5	40.5	569	9	ACU13033	Acu13033 DNA clone
19	486	35.9	651	9	ACU13113	Acu13113 DNA clone
20	461	34.1	675	9	ACU13102	Acu13102 DNA clone
21	447	33.0	974	12	ADM47592	Adm47592 Polynucle
22	447	33.0	990	3	AAZ37819	Aaz37819 Arabidops
23	429	31.7	1163	12	ADM47593	Adm47593 Polynucle
24	417	30.8	522	9	ACU13056	Acu13056 DNA clone
25	417	30.8	599	9	ACU13064	Acu13064 DNA clone
26	417	30.8	602	9	ACU13077	Acu13077 DNA clone
27	417	30.8	634	9	ACU13076	Acu13076 DNA clone
28	407	30.1	582	9	ACU13025	Acu13025 DNA clone
29	406.5	30.0	555	9	ACU13060	Acu13060 DNA clone
30	406.5	30.0	603	9	ACU13078	Acu13078 DNA clone
31	406.5	30.0	625	9	ACU13062	Acu13062 DNA clone
32	406.5	30.0	632	9	ACU13080	Acu13080 DNA clone
33	406.5	30.0	656	9	ACU13087	Acu13087 DNA clone
34	405.5	30.0	661	9	ACU13030	Acu13030 DNA clone
35	405.5	30.0	668	9	ACU13022	Acu13022 DNA clone
36	403.5	29.8	933	3	AAZ50752	Aaz50752 Arabidops
37	401.5	29.7	439	9	ACU13058	Acu13058 DNA clone
38	401.5	29.7	551	9	ACU13082	Acu13082 DNA clone
39	401.5	29.7	623	9	ACU13068	Acu13068 DNA clone
40	401.5	29.7	638	9	ACU13086	Acu13086 DNA clone
41	401.5	29.7	639	9	ACU13098	Acu13098 DNA clone
42	401.5	29.7	644	9	ACU13084	Acu13084 DNA clone
43	401.5	29.7	650	9	ACU13097	Acu13097 DNA clone
44	400.5	29.6	673	9	ACU13032	Acu13032 DNA clone
45	400.5	29.6	937	3	AAZ38614	Aaz38614 Arabidops

#### ALIGNMENTS

RESULT 1	AAZ35493	
ID	AAZ35493	standard; cDNA; 1205 BP.
XX		
AC	AAZ35493;	
XX		
DT	15-SEP-2003	(revised)
DT	11-APR-2000	(first entry)
XX		
DE	Petunia nectary-specific NECL cDNA.	
XX		
KM	NECL; nectary; nectar; transgenic plant; honey; ds.	
XX		
OS	Petunia x hybrida.	
XX		
FH	Key	Location/Qualifiers
FT	CDS	79..876
FT		/*tag= a
XX		
PN	EP974667-A1.	
XX		
PD	26-JAN-2000.	
XX		
PF	16-JUL-1998;	98EP-00202375.
XX		
PR	16-JUL-1998;	98EP-00202375.
XX		
PA	(CPRO-) CPRO-DLO CENT PLANTENVEREDELINGS REPROD.	
XX		
PI	Creemers J, Angenent GC, Kater MM;	

XX DR WPI; 2000-108400/10.  
 XX P-PSDB; AAY58647.  
 XX PT Novel DNA sequences used to produce modified honey, the metabolites of  
 XX PT which can be isolated and purified.  
 XX PS Claim 2; Page 16; 56pp; English.  
 XX CC The present sequence is that of Petunia hybrida strain W15 NECL cDNA, as  
 CC produced from 2 overlapping partial clones (see AA235497-98) obtained by  
 CC differential display RT-PCR and RACE PCR. The NECL gene encodes a 265-  
 CC amino acid protein (see AAY58647). NECL is highly expressed in the  
 CC nectaries of petunia and weakly expressed in the stem. The present  
 CC invention provides a method for producing recombinant proteins in honey.  
 CC The honey is manufactured by insects, preferably honeybees, that collect  
 CC the nectar of transgenic plants. The NECL gene and its promoter (see  
 CC AA235496) can be utilized in expression cassettes for the production of  
 CC transgenic plants that produce a protein of interest in their nectar. The  
 CC function of NECL has not yet been determined. (Updated on 15-SEP-2003 to  
 CC standardise OS field)  
 XX SQ Sequence 1205 BP; 374 A; 219 C; 229 G; 383 T; 0 U; 0 Other;  
 Alignment Scores:  
 Pred. No.: 6.8e-143 Length: 1205  
 Score: 1353.00 Matches: 265  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 3 Gaps: 0  
 US-09-743-885A-1 (1-265) x AA235493 (1-1205)  
 QY 1 MetAlaGlnLeuAArgAlaAspAspLeuSerPheIlePheGlyLeuGlyAsnIleVal 20  
 Db 79 ATGGCGCAATTCAGTCTGCTGATGACTTGTCTTTCATATTTGGCTTCTTGGAATATTTGTA 138  
 QY 21 SerPheMetValPheLeuAlaProValProThrPheTyrLysIleTyrLysArgLysSer 40  
 Db 139 TCATTCAATGGTCTCTTACAGACCCGTGCACAACATTTTACAAATATATATAAGAAATCA 198  
 QY 41 SerGlnGlyTyrGlnAlaIleProTyrMetValAlaLeuPheSerAlaGlyLeuLeuLeu 60  
 Db 199 TCAGAAAGAAATCAAGCAATACCATATATGTTACACTGTTCAAGCGCGCACTATTGCTA 258  
 QY 61 TyrTyrAlaTyrLeuAArgLysAsnAlaTyrLeuIleValSerIleAsnGlyPheGlyCys 80  
 Db 259 TATTATGCTTATCTCAGAGAGAAATGCTTATCTTATCGTCAGCATTAATGGCTTGGATGT 318  
 QY 81 AlaIleGlnLeuThrTyrIleSerLeuPheLeuPheTyrAlaProArgLysSerLysIle 100  
 Db 319 GCCATTGAATTAACATATATCTCTCTGTTCTCTTTACGGGCCCAAAAGTCTAAAGATT 378  
 QY 101 PheThrGlyTyrPheMetLeuLeuGlnLeuGlyAlaLeuGlyMetValMetProIleThr 120  
 Db 379 TTCAACAGGGGCGCGATGCTCTTGAATGGGACCCCTAGGAATGATGATGCCAATTACT 438  
 QY 121 TyrLeuLeuAlaGlnGlySerPheIleArgValMetIleValGlyTyrPheCysAlaAlaIle 140  
 Db 439 TATTATTAGAGAGAGCTCACTACATAGATGATGATGATGGATGGATTTGTGCGCTATC 498  
 QY 141 AsnValAlaValPheAlaAlaProLeuSerIleMetArgIleValIleLysThrLysSer 160  
 Db 499 AATGTTCTGCTTGTGCTGCTCTTTAAGCATCATAGGCAAGTAATATAAACAAGT 558  
 QY 161 ValGlnPheMetProPheThrLeuSerLeuPheLeuThrLeuCysAlaIleMetTyrPhe 180  
 Db 559 GTAGAGTTCAAGCCCTTCACTTATCTTTGTTCTCTGCTGTCACATATGCTTT 618  
 QY 181 PheTyrGlyPhePheLysLysAspPheTyrIleAlaPheProAsnIleLeuGlyPheLeu 200  
 Db 619 TTCTATGGGTTTTCAGAGAGGACTTTTACATTCGTTTCCAAATATATACGGGCTTCTA 678

QY 201 PheGlyIleValGlnMetLeuThrPheValTyrLysAspSerLysArgIleAspAsp 220  
 Db 679 TTCGGAATGCTTCAATGCTATATATTTGTTTACAGAGATTCAAAGAAATAGATGAT 728  
 QY 221 GlnLysSerAspProValArgGlnAlaThrLysSerLysGlnGlyValGlnIleIle 240  
 Db 739 GAAAAATCTGATCTCTTCGAGAGGCTACAAATCAAAAGAGGTGTAGAAATCATTTATC 798  
 QY 241 AsnIIGlnAspAspAsnSerAspAsnAlaLeuGlnSerMetGlnLysAspPheSerArg 260  
 Db 799 AACATTGAAGATATATATTTCTGATTAACGATTCAGTCCATGAGAGAAATTTTCCAGA 858  
 QY 261 LeuArgThrSerLys 265  
 Db 859 CTCGAGACATCAAAA 873  
 RESULT 2  
 AA250200  
 ID AA250200 standard; cDNA; 1205 BP.  
 XX AC AA250200;  
 XX DT 15-SEP-2003 (revised)  
 DT 17-MAY-2000 (first entry)  
 DE Petunia hybrida nectary-specific protein NECL cDNA.  
 DE NECL; nectary-specific protein; metabolite; recombinant protein;  
 KW transgenic plant; nectar; honey; insect; bee; pharmaceutical; enzyme;  
 KW biotect; antioxidant; food additive; ss.  
 XX OS Petunia x hybrida.  
 XX FH Key Location/Qualifiers  
 FT CDS 79..876  
 FT /tag=a  
 FT /product="NECL protein"  
 FT  
 PN MO200004176-A1.  
 XX PD 27-JAN-2000.  
 PD 15-JUL-1999; 99MO-NL000453.  
 PF 16-JUL-1998; 98BP-00202375.  
 PR 14-DEC-1998; 98BP-00204215.  
 XX PA (CPRO-) CPRO-DIO CENT PLANTENVERDELINGS REPROD.  
 XX PI Creemers J, Angenent GC, Kater MM;  
 XX WPI; 2000-182438/16.  
 DR P-PSDB; AAY44803.  
 DR  
 XX New DNA encoding Petunia hybrida nectary-specific proteins, useful for,  
 PT e.g. producing modified honey.  
 PT  
 XX  
 PS Claim 8; Page 42; 93pp; English.  
 XX CC The present sequence is a cDNA encoding nectary-specific protein NECL. It  
 CC was isolated from nectaries of Petunia hybrida strain W15 using mRNA  
 CC Differential Display system. NECL resembles membrane protein and is  
 CC strongly expressed in the nectaries of P. hybrida. A DNA sequence from  
 CC the promoter region upstream of nectary-specific expressed sequence e.g.  
 CC NECL and FBP15 DNAs is used in a recombinant DNA construct comprising a  
 CC DNA encoding a metabolite preferably recombinant protein, a DNA encoding  
 CC a signal peptide that targets the recombinant protein to the nectar and  
 CC optionally a signal sequence functional in plants for the transcription  
 CC termination and polyadenylation of an RNA molecule. The DNA construct is  
 CC useful for producing transgenic plants which excrete recombinant proteins  
 CC in its nectar. The nectar is processed into honey by insects (preferably  
 CC bees) and the desired protein is easily recovered from it. The



```
QY 81 AlallegluethrTyrIleSerIeuPheUeUyralaProArgLySerIle 100
Db 319 GCCATTGAATTAACATATATCTCTGTTCTTTTACGGCCGACGAAGATTAAGATT 378
QY 101 PheThrGlyTTPleuMetIeuGluLeuGlyAlaIeuGlyMetValMetProIleThr 120
Db 379 TTCACAGGGTGGCTGATGCTCTTAGAATTGGAGCCCTTAGAAGATGGATGCCAATTACT 438
QY 121 TyrIeuLeuAlaGluGlySerHisArgValMetIleValGlyTTPleCysAlaAlaIle 140
Db 439 TATTATTAGCAGAAAGCTCATAGACTGATGATAGTGGATGGATTGGACGCTATC 498
QY 141 AsnValAlaValPheAlaAlaProIeuSerIleMetArgGlnValIleYsThrLySer 160
Db 499 AATGTGCTGCTTGTGGCTCTTTAAGCATCATGAGGCAAGTAAATAAACAAGAGT 558
QY 161 ValGluPheMetProPheThrIeuSerIeuPheUeUyralaProArgLySerIle 180
Db 559 GTAGAGTTCATGCTTCACTTATCTTGTCTTCACTCTGTGCGCATATGTGGT 618
QY 181 PheTyrGlyPhePheLyAspPheTyrIleAlaPheProAlaIleLeuGlyPheUeU 200
Db 619 TTCATAGGCTTTTCAAGAGGACTTTTACATTCGCTTCCAAATATATCGGCTTTCTA 678
QY 201 PheGlyIleValGlnMetIeuLeuTyrPheValTyrLyAspSerIleArgIleAspAsp 220
Db 679 TTCGAATCGTTCAATGCTATATATTTGTTTACAAAGATTCAAAGAAATGATGAT 738
QY 221 GluLySerAspProValArgGlnAlaThrLySerIleGlyValGlnIleIle 240
Db 739 GAAATAATCTGATCTGTTCGAAAGCTACAAATAAGAGGTGTAAGAAATCATATATC 798
QY 241 AsnIleGluAspAspAsnSerAspAsnAlaIeuGlnSerMetGluLyS 256
Db 799 AACATTGAAGATGATATCTGATACCGCATTCGATCGAGGAAG 846

RESULT 4
AAZ50207
ID AAZ50207 standard; DNA; 847 BP.
XX
AC AAZ50207;
XX
DT 15-SEP-2003 (revised)
DT 17-MAY-2000 (first entry)
XX
DE Clone RC8 for cloning P. hybrida NEC1 gene.
XX
KW NEC1; nectary-specific protein; metabolite; recombinant protein;
KW transgenic plant; nectar; honey; insect; bee; pharmaceutical; enzyme;
KW biotect; antioxidant; food additive; clone RC8; ds.
XX
OS Petunia x hybrida.
XX
FH Key Location/Qualifiers
FT primer_bind complement(79..100)
FT /*tag= a
FT /bound_moiety= "Prat 129 primer"
FT /note= "Prat 129 is used with primer prat 122 to amplify
FT the coding region of NEC1 cDNA"
XX
XX WO200004176-A1.
XX
XX 27-JAN-2000.
XX
XX 15-JUL-1999; 99WO-NL000453.
XX
XX 16-JUL-1998; 98EP-00202375.
XX
XX 14-DEC-1998; 98EP-00204215.
XX
XX (CPRO-) CPRO-DLO CENT PLANTENVERDELINGS REPROD.
XX
XX Creemers J, Angenent GC, Kater MM;
PI
```

```
XX
DR WPI; 2000-182438/16.
XX
PT New DNA encoding Petunia hybrida nectary-specific proteins, useful for,
PT e.g. producing modified honey.
XX
XX Example 1; Fig 3; 93pp; English.
XX
CC The patent discloses a recombinant DNA construct comprising a DNA
CC sequence from the promoter region upstream of nectary-specific expressed
CC sequence e.g. NEC1 and FRP15 DNAs, a DNA encoding a metabolite preferably
CC recombinant protein, a DNA encoding a signal peptide that targets the
CC recombinant protein to the nectar and optionally a signal sequence
CC polyaenylation of an RNA molecule. The DNA construct is useful for
CC producing transgenic plants which excrete recombinant proteins in its
CC nectar. The nectar is processed into honey by insects (preferably bees)
CC and the desired protein is easily recovered from it. The recombinant
CC proteins are useful for pharmaceutical purposes, as enzymes for biotests
CC and antioxidants for food additives. The present DNA sequence is that of
CC clone RC8 obtained by 5' RACE PCR of cDNA derived from nectaries of P.
CC hybrida flowers. The overlapping sequences of this clone and a 3' cDNA
CC clone DP18a were used to isolate the full length cDNA of P. hybrida NEC1
CC gene which is strongly expressed in nectaries. (updated on 15-SEP-2003 to
CC standardise OS field)
XX
SQ Sequence 847 BP; 249 A; 157 C; 172 G; 269 T; 0 U; 0 Other;
XX

Alignment Scores:
Pred. No: 3.95e-138 Length: 847
Score: 1309.00 Matches: 256
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 96.75% Indels: 0
DB: 3 Gaps: 0

US-09-743-885A-1 (1-265) x AAZ50207 (1-847)
QY 1 MetAlaGlnIeuArgIlaIaPaspIeuSerPheIlePheGlyIeuLeuGlyAsnIleVal 20
Db 79 ATGGCGCAATTAGCTGCTGATGATGCTTCTTATATTTGGCTTCTGATATATTGTA 138
QY 21 SerPheMetValPheUeUyralaProValProThrPheTyrLySileTyrLySArgLySer 40
Db 139 TCATTCATGCTTCTTCCAGCACCCGCGCAACATTTTAAACAATAATATAAGGAATCA 198
QY 41 SerGluGlyTyrGlnAlaIleProTyrMetValAlaIeuPheSerIleGlyIeuLeu 60
Db 199 TCGAAGAGATATACAGCAATATCATATATGTGACGTTCAGCGCGGACTATATGCTA 258
QY 61 TyrTyrAlaTyrIeuArgLyAsnAlaTyrIleUleValSerIleAsnGlyPheGlyCys 80
Db 259 TATTATGCTTATCTCAGAAAGATGCTTATCTGATGCTGATTAATGCTTTGATGAT 318
QY 81 AlallegluethrTyrIleSerIeuPheUeUyralaProArgLySerIle 100
Db 319 GCCATTGAATTAACATATATCTCTGTTCTTTTACGGCCGACGAAGTAAAGTAAAGATT 378
QY 101 PheThrGlyTTPleuMetIeuGluLeuGlyAlaIeuGlyMetValMetProIleThr 120
Db 379 TTCACAGGGTGGCTGATGCTCTTAGAATTGGAGCCCTTAGAAGATGGATGCCAATTACT 438
QY 121 TyrIeuLeuAlaGluGlySerHisArgValMetIleValGlyTTPleCysAlaAlaIle 140
Db 439 TATTATTAGCAGAAAGCTCATAGACTGATGATAGTGGATGGATTGGACGCTATC 498
QY 141 AsnValAlaValPheAlaAlaProIeuSerIleMetArgGlnValIleYsThrLySer 160
Db 499 AATGTGCTGCTTGTGGCTCTTTAAGCATCATGAGGCAAGTAAATAAACAAGAGT 558
QY 161 ValGluPheMetProPheThrIeuSerIeuPheUeUyralaProArgLySerIle 180
Db 559 GTAGAGTTCATGCTTCACTTATCTTGTCTTCACTCTGTGCGCATATGTGGT 618
```



QY 181 PheTYGlyPhePheLYSAspPheTYrIleAlaPheProXenileuGLyPheLeu 200  
Db 619 TTCATAGGGTTTTCAGAAAGACCTTTACATTCGCTTCCAATATACCTGGCTTCTA 678  
QY 201 PheGlyIleValGlnMetLeuLeuTYrPheValTYrLYSAspSerLYSArGIIeAspAsp 220  
Db 679 TTCGAATCGTTCAAATGCTATTATATTTGTTTACAGAGATTCAAAGAGATAGATGAT 728  
QY 221 GlnLYSAspProValArgGlnAlaThrLYSAspSerLYSGLyValGlnIleIleIle 240  
Db 739 GAAATAATCGATCCTGTCAGAGAGCTACAAATCAAAAGAGCTGTAAGATCATTATC 798  
QY 241 AenIIeGlnAspAspAsnSerAspAsnAlaLeuGlnSerMetGlnLYS 256  
Db 799 AACATTGAAGATGATTAATTTGATTAACGCATTCGATCCATCGAGAG 846  
RESULT 5  
AAC47558  
ID AAC47558 standard; DNA; 900 BP.  
XX AAC47558;  
AC  
XX 18-OCT-2000 (first entry)  
DT  
XX  
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 54267.  
XX  
KM Hybridisation assay; genetic mapping; gene expression control;  
KM protein identification; signal transduction pathway; metabolic pathway;  
KM promoter; termination sequence; ss.  
XX  
OS Arabidopsis thaliana.  
PN EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-00301439.  
XX  
PR 25-FEB-1999; 99US-0121825P.  
PR 05-MAR-1999; 99US-0123180P.  
PR 09-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 23-APR-1999; 99US-0130891P.  
PR 28-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 30-APR-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 06-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.

PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140053P.  
PR 23-JUN-1999; 99US-0140054P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145313P.  
PR 27-JUL-1999; 99US-0145318P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.



KM Hybridization assay; Genetic mapping; Gene expression control;  
KM protein identification; signal transduction pathway; metabolic pathway;  
XX promoter, termination sequence; ss.  
OS Arabidopsis thaliana.  
PN EP1033405-A2.  
XX  
PD 06-SEP-2000.  
PF  
XX 25-FEB-2000; 2000EP-00301439.  
XX  
PR 25-FEB-1999; 99US-0121825P.  
PR 05-MAR-1999; 99US-0123180P.  
PR 09-MAR-1999; 99US-0123568P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 28-APR-1999; 99US-0130691P.  
PR 30-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 04-MAY-1999; 99US-0132407P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134321P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.

PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 05-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 06-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 23-AUG-1999; 99US-0150566P.  
PR 25-AUG-1999; 99US-0150884P.  
PR 26-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.

```
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 26-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-015753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 23-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.
PR 29-OCT-1999; 99US-0162142P.
```

## Alignment Scores:

```
Pred. No.: 1.23e-69 Length: 902
Score: 703.50 Matches: 136
Percent Similarity: 74.49% Conservaive: 48
Best Local Similarity: 55.06% Mismatches: 62
Query Match: 52.00% Indels: 1
Gaps: 3
```

US-09-743-885A-1 (1-265) x AAC40321 (1-902)

```
QY 4 LeuThrTyrIleSerLeuPheLeuTyrAlaProArgIysSerIysIlePheThrGly 23
Db 51 CTCGAGGTTCAATGAAATGCTTTCTCTTGGCTCTTGAACAATGTCCTTTGGG 110
QY 24 ValPheLeuAlaProValProThrPheTyrIleTyrIleTyrIleTyrIleTyrIle 43
Db 111 GTGTTCTTGTGACAGTCACAGTCGACGTTTATGCAATATACAGAAATATCAAAAGG 170
QY 44 TyrGlnAlaIlePheTyrMetValAlaLeuPheSerIleGlyLeuLeuTyrTyrAla 63
Db 171 TTTCAGTCGATACCGTCATATGTCACATACCAAGTCACTTCTTCTTACTACGCA 230
QY 64 TyrLeuArgIysAenAlaTyrLeuIleValSerIleAngIyPheGlyCyAAlaIleGlu 83
Db 231 ATCTGAAGACACATGCTTATCTATCTATTAATTAACACCTTGGATGTTCAATGAA 290
```

```
QY 84 LeuThrTyrIleSerLeuPheLeuTyrAlaProArgIysSerIysIlePheThrGly 103
Db 291 ATCTCCACTTATGTTTCTATCTATCTTATGACCAAGAGGCGCAAGATATCCACGTTG 350
QY 104 TrpLeuMetLeu---GluLeuGlyAlaLeuGlyMetValMetProIleThrTyrLeu 122
Db 351 AAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 410
QY 123 LeuAlaGluIlySerIleArgValMetIleValGlyTyrIleCyAAlaIleAenVal 142
Db 411 TTGTTCCAAACAAACACCCGCTGACCGTTGATGATGATGATGATGATGATGATGATGAT 470
QY 143 AlaValPheAlaAlaProIleSerIleMetArgGlnValIleTyrThrIlySerValGlu 162
Db 471 GCAGTCTTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 530
QY 163 PheMetProPheThrLeuSerLeuPheLeuThrLeuCyAAlaIleMetTrpPhePheTyr 182
Db 531 TACATGCCGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTAT 590
QY 183 GlyPhePheIlySerIlySerPheTyrIleAlaPheProAenIleLeuGlyPheLeuPheGly 202
Db 591 GCACTTCTTATCAAAAGCAAGTCATTCATGCAAAACATTCGCTTTTCTATTCGCT 650
QY 203 IleValGlnMetLeuLeuTyrPheValTyrIlyApsSerIlyAenIleApsApsGluIys 222
Db 651 GTAGCTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 710
QY 223 SerApsProValArgGluAlaThrIlySerIlyGluGlyAlaGluIleIleAenIle 242
Db 711 GAGAACCACTTACCAAAATTAACCGATGTTAAGAAAGTCCGATCGGCCGTTGAATG 770
QY 243 GluApsApsAenSerApsAen 249
Db 771 CCGATGTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 791
RESULT 7
AB213458 standard; DNA; 879 BP.
ID AB213458
XX
AC AB213458;
XX
DT 21-JAN-2003 (first entry)
XX
DE Arabidopsis thaliana stress regulated gene SEQ ID NO 1263.
XX
KW Arabidopsis thaliana; plant; gene; stress; transgenic; ds.
XX
OS Arabidopsis thaliana.
XX
PN WO200216655-A2.
XX
PD 28-FEB-2002.
XX
PF 24-AUG-2001; 2001MO-US026685.
XX
PR 24-AUG-2000; 2000US-0227866P.
PR 26-JAN-2001; 2001US-0264647P.
PR 22-JUN-2001; 2001US-0300111P.
XX
PA (SGRI ) SCRIPPS RES INST.
PA (SYGN ) SYNGENTA PARTICIPATIONS AG.
XX
PI Harper JF, Krebs J, Wang X, Zhu T;
XX
DR WPI; 2002-304127/34.
XX
PT Identifying a stress condition to which a plant cell has been exposed and
XX producing plants with increased tolerance to these abiotic stresses.
XX
PS Claim 15; SEQ ID NO 1263; 577pp + Sequence Listing; English.
XX
CC The invention relates to identifying a stress condition to which a plant
```

cell has been exposed, comprising: (a) contacting nucleic acid representative of expressed polynucleotides in the plant cell with an array or probes representative of the plant cell genome; and (b) detecting a profile of expressed polynucleotides in the plant cell characteristic of a stress response. The method is useful in the production of transgenic plants, cells and seeds and in producing plants with increased tolerance to abiotic stresses. The present sequence is that of an Arabidopsis thaliana stress regulated gene (AB212196-AB217574) used in methods of the invention. Note: The sequence data for this patent is not represented in the printed specification but is based on sequence information supplied to Derwent by the European Patent Office

CC Sequence 879 BP; 215 A; 199 C; 195 G; 270 T; 0 U; 0 Other;

#### Alignment Scores:

Pred. No.:	1,31e-60	Length:	879
Score:	623.50	Matches:	119
Percent Similarity:	65.31%	Conservative:	58
Best Local Similarity:	43.91%	Mismatches:	63
Query Match:	46.08%	Indels:	31
DB:	6	Gaps:	4

US-09-743-885a-1 (1-265) x AB213458 (1-879)

```
Qy 9 LeuSerPheIlePheGlyLeuLeuGlyAenIleValSerPheMetValPheLeuAlaPro 28
Db 28 CTGCGTTTATCTTCGCGATCTTGAAGAAAGCTATATCTTCTGTTATCTTCGCTCCA 87
Qy 29 ValProThPheTyrlYsIleTyrlYsArgLysSerSerGluGlyTyrGlnAlaIlePro 48
Db 88 GTGCCAACTTTTATAGATATACAGAGAAATTCAGCGAAAGTTCCAGTCGCTACCG 147
Qy 49 TyrMetValAlaLeuPheSerAlaGlyLeuLeuLeuTyrlYsAlaTyrLeuAlaGlyLeu 68
Db 148 TACCAAGTGCCTGATTTAGCTCAGTCATGCTCTACACCATGATGATGAAGAAGAC 207
Qy 69 AlaTyrLeuIleValSerIleAenGlyPheGlyCysAlaIleGluLeuThyrlYsIleSer 88
Db 208 GCTTTCTCTTATTTACCATCACTCTCTTGGCTGCTGCTGAGACTCTCTACATAGCC 267
Qy 89 LeuPheLeuPheTyrlYsAlaProArgLysSerLysIlePheThyrlYsIleLeuMetLeu 107
Db 268 ATGTTCTTGCGCTTACCGCACGAGGAGAAAGATTCGCGCTATGAAAGTTGTTCAATGCA 327
Qy 108 LeuGluLeuGlyAlaLeuGlyMetValMetProIleThyrlYsIleLeuAlaGlySer 127
Db 328 ATGAAAGTTCCTTCTTCGTTGATTTCTATGATGATACATATTCGTGTTAAACTCT 387
Qy 128 Hle---ArgValMetIleValGlyTyrPheCysAlaAlaIleAenValAlaValPheAla 146
Db 388 CCCCTCCAAAGTCTCTGACTCGGCTGATTTGTTGCTTCATTTCTGTTTCGCT 447
Qy 147 AlaProLeuSerIleMetArgIleValIleLysThyrlYsSerValGluPheMetProPhe 166
Db 448 GCCCTCTATATGATCGGCTCGTGATTAAGCAAAAGGTGGAGTGCATGCCCTTC 507
Qy 167 ThrLeuSerLeuPheLeuThyrlYsCysAlaThMetTyrPhePheTyrlYsPhePheLys 186
Db 508 ACCCTTCTTCTTCCTCACTATAGCGCCGTTATGATGTTCCCTTATGTTATTCCTC 567
Qy 187 LysAspPheTyrlYsAlaPheProAenIleLeuGlyPheLeuPheGlyIleValGlnMet 206
Db 568 AATGACATATGATGATGATTCCAACGTTGGGATTCGTTACTAGGCGCTTGAACATG 627
Qy 207 LeuLeuTyrlYsValTyrlYsAspSerLysArg----- 217
Db 628 GTTTTACTGTTTATACAGAACTCAATAGAGAAACCAAGAAAGATTATGTCAGAA 687
Qy 218 -----Ile 218
Db 688 CAACAACTTAAAGATATTGTCGATGATGTCGCTTATGATGTCGAAAGTGCACCCAGTT 747
Qy 219 AspAspGluLysSerAspProValArgGluAlaThyrlYsSerLysGluGlyValGluIle 238
```

```
Db 748 GTGACGGAATCGGTGAGCCACCATCTCTGAAGCGTT---CATCATGAGATCTGTCCAA 804
Qy 239 IleIleAenIleGluPheAspAsnSerAspAsn 249
Db 805 GTTACTTAAGTGGAGGAGCGCTCAATTGAAAC 837
RESULT 8
AAC33408
ID AAC33408 standard; DNB; 1263 BP.
AC AAC33408;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 2927.
XX
XX Hybridisation assay; genetic mapping; gene expression control;
KM protein identification; signal transduction pathway; metabolic pathway;
KM promoter; termination sequence; ss.
XX
XX Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
XX 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 22-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135313P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138547P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-01391199P.
PR 16-JUN-1999; 99US-0139452P.
```

PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139839P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143622P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144637P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146387P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147302P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.

PR 13-AUG-1999; 99US-0148665P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156569P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158022P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161922P.  
PR 28-OCT-1999; 99US-0161933P.  
PR 29-OCT-1999; 99US-0162142P.

## Alignment Scores:

Pred. No.: 2,16e-60 Length: 1263  
Score: 623.50 Matches: 119  
Percent Similarity: 65.31% Conservative: 58  
Best Local Similarity: 43.91% Mismatches: 63  
Query Match: 46.08% Indels: 31  
DB: 3 Gaps: 4

US-09-743-885A-1 (1-265) x AAC33408 (1-1263)

```

QY 9 LeuSerPheIlePheGlyLeuLeuGlyAenIleValSerPheMetValPheLeuAlaPro 28
DB 162 CTGCGTTTATCTTGGCATCTTGAAGAAAGTGATCTCTTCTGATTCCTGGCTCCA 221
QY 29 ValProThPheTyrlsIleTyrlsArgLySerSerGluGlyTyrglnAlaIlePro 48
DB 222 GTGCCAACTTTTATGAAATATACAGAGAAATCCAGCGAAAGTTTCCAGTGGCTACCG 281
QY 49 TyrMetValAlaLeuPheSerAlaGlyLeuLeuLeuTyrlsAlaTyrlsLeuArgLyAsn 68
DB 282 TACCAAGTGTGGTATTTAGCTGCATGCTATGCTTACTACCACTTGAATTAAGAAAGAC 341
QY 69 AlaTyrlsIleValSerIleAsnGlyPheGlyCysAlaIleGluLeuThyrlsIleSer 88
DB 342 GCTTTCTCTTAATTAACATCACTCTTGGCTGGCTGGAGACTCTCTACATAGCC 401
QY 89 LeuPheLeuPheTyrlsAlaProArgLySerLyIlePheThyrlsTyrlsLeuMetLeu 107
DB 402 ATGTTCTTGGCTTACGCCACACGAGAAAGATTCGGCTTGAAGTTGTTTCATAGCA 461
QY 108 LeuGluLeuGlyAlaLeuGlyMetValMetProIleThyrlsLeuLeuAlaGlyIleSer 127
DB 462 ATGAACGTTGGCTTCTCTGCTGATTTCTAATGTAACATTTCTGTTAAACTCCT 521
QY 128 His---ArgValMetIleValGlyTyrPileCysAlaAlaIleAsnValAlaValPheAla 146
DB 522 CCCCTCCAACTCTCTACTCGCTGAGATTTGTTGGCCATTTCTGTTCTGTTTCGCT 581
QY 147 AlaProLeuSerIleMetArgIleValIleTyrlsTyrlsSerValGluPheMetProPhe 166
DB 582 GCCCTCTAATATATCGTGGCTGCTGATTAACAAAGAGTGTGAGTACATGCCCTTC 641
QY 167 ThrLeuSerLeuPheLeuThyrlsCysAlaIleMetTrpPheThyrlsGlyPhePheLyS 186
DB 642 ACGCTTCTTCTTCTCTCTCACTAATACGCCGCTTATGTGTTCCCTTATGTTTATTCCTC 701
QY 187 LyAsnProPheTyrlsIleAlaPheProAlaIleGlyPheLeuPheGlyIleValIleMet 206
DB 702 AATGACATATGCATACCGATTCCTCAACGCTGAGTGTGATCTACTAGGCGCTTGCAATG 761
QY 207 LeuLeuTyrlsPheValTyrlsAsnSerLyAsnArg----- 217
DB 762 GTTTTCTACTGTGTTTACAGAACTCAATGAGAAACAGAGAAATTAATTCGTAGAA 821
QY 218 -----11e 218
DB 822 GAACAATTAGAGTATTTGTCGATGATGCTCGTGTAGTGTGCGAAGTGCACCCAGTT 881
QY 219 AsnAspGluLySerAsnProValArgGluAlaThyrlsSerLySerGluGlyValGluIle 238
DB 882 GTGACGGAATTCGTGAGCCACTCTCTGAAGCCGTT--CATCATGAGATGTGCCAAA 938
QY 239 IleIleAsnIleGluAsnAsnSerAsn 249
DB 939 GTTACTAAGTGGAGAGAGCCGCTCAATTGAAGAAC 971

```

```

PN W02004035798-A2.
XX 29-APR-2004.
XX 20-OCT-2003; 2003WO-EP011658.
XX 18-OCT-2002; 2002EP-00079408.
XX (CROP-) CROPDSEIGN NV.
XX Inze D, De Veylder L, Vlieghe K;
XX P-PSDB; ADN73901.
XX WPI; 2004-348466/32.
XX P-PSDB; ADN73901.
XX Altering plant characteristics, useful for producing plants for enzyme or
XX pharmaceutical production comprises modifying in a plant, expression of
XX one or more nucleic acids and/or modifying level or activity of one or
XX more proteins.
PS Claim 1, SEQ ID NO 1795; 134pp; English.
XX
XX This invention relates to a novel method for altering one or more plant
XX characteristics. Specifically, it refers to identifying genes that are up
XX - or down-regulated in transgenic plants overexpressing the heterodimeric
XX E2Fa/Dpa transcription factor of Arabidopsis and using these sequences to
XX alter plant characteristics accordingly. The present invention describes
XX generating transgenic plants for the production of growth regulators,
XX enzymes, therapeutic, pharmaceuticals and animal feed products, where
XX the altered plant characteristics are selected from increased yield or
XX biomass, enhanced survival capacity, stress tolerance, plant architecture
XX or physiology, altered endoreplication, biochemistry, signal
XX transduction, storage lipid mobilisation and/or altered photosynthesis,
XX each relative to the corresponding wild type plants. Accordingly, these
XX sequences can also be useful as positive or negative selectable markers
XX during transformation of cells or tissues. The identified genes play a
XX role in a variety of biological processes such as DNA replication, cell
XX wall biosynthesis, nitrogen and/or carbon metabolism or they function as
XX transcription factors. This polynucleotide sequence is thale cress cDNA
XX repressed 1.3 fold or more in plants overexpressing the E2Fa/Dpa
XX transcription factor, given in an exemplification of the invention.
XX
XX Sequence 870 BP; 210 A; 209 C; 193 G; 258 T; 0 U; 0 Other;
XX
XX Alignment Scores:
XX Pred. No.: 1,91e-60 Length: 870
XX Score: 622.00 Matches: 126
XX Percent Similarity: 65.40% Conservative: 46
XX Best Local Similarity: 47.91% Mismatches: 63
XX Query Match: 45.97% Indels: 28
XX DB: 12 Gaps: 4
XX
XX US-09-743-885A-1 (1-265) x ADN73900 (1-870)
QY 10 SerPheIlePheGlyLeuLeuGlyAenIleValSerPheMetValPheLeuAlaProVal 29
DB 31 GCGTTTGTCTTGGCTTCTCGCAACTTATCTCTTCTGCTGCTGCTGCTGCTGCTG 90
QY 30 ProThPheTyrlsIleTyrlsArgLySerSerGluGlyTyrglnAlaIleProTy 49
DB 91 CCAAGCTTCTAATGAGATTTGAAGAAAGACAGACAGAAAGGTTTCAGTCTTCTTAT 150
QY 50 MetValAlaLeuPheSerAlaGlyLeuLeuLeuTyrlsAlaTyrlsLeuArgLyAsnAla 69
DB 151 GTTGTGGCGCTCTTCAAGTGCAGCGCTTGTCTTACTAAGCAGACAGAAAGAAAGTGT 210
QY 70 TyrLeuIleValSerIleAsnGlyPheGlyCysAlaIleGluLeuThyrlsIleSer 89
DB 211 TTCTCTCTGTAACCTTACCGCTTGTGCTTCTCATGAAACATCTACATCTTATG 270
QY PheLeuPheTyrlsAlaProArgLySerLyIlePheThr---GlyTyrlsLeuMetLeu 108
DB 271 TTCTTGTCTTACGCTCCCAAGCGAGCTCGAGTGTGACAGTGAAGATGCTACTTATG 330

```





```

PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145313P.
PR 27-JUL-1999; 99US-0145318P.
PR 27-JUL-1999; 99US-0145319P.
PR 28-JUL-1999; 99US-0145351P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 11-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149802P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-015138P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.

```

```

PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161040P.
PR 25-OCT-1999; 99US-0161045P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 26-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Alignment Scores:
Pred. No.: 2,89e-60 Length: 1177
Score: 622.00 Matches: 126
Percent Similarity: 65.40% Conservative: 46
Best Local Similarity: 47.91% Mismatches: 63
Query Match: 45.97% Indels: 28
DB: 3 Gaps: 4

US-09-743-885a-1 (1-265) x AAC51633 (1-1177)

QY 10 SerpPhlePheGlyLeuLeuGlyValSerPheMetValPheLeuAProVal 29
DB 115 GCCTTGTCTTGGCTTGCAGCACTTATCTCTTCCGCTGCTTCTATCTCTG 174

QY 30 ProthPheTyrlleValSerPheGlySerSerGlyTyrlleValleProTy 49
DB 175 CCACGCTTATAGATTGGAAGAAAGACACAGAGGTTTCAGTCTATCTTAT 234

QY 50 MetValAlaLeuPheSerAlaGlyLeuLeuTyrlleValleTyrlleValle 69
DB 235 GTTGGCCGCTTTCAGTCCGAGCGCTTGGCTTCTTACAGACAGAAAGATGTC 294

QY 70 TyrlleValleValSerPheGlySerPheGlyValleGlyLeuThTyrlleSer 89
DB 295 TTCTCTCTGTAACCTTAAACCTTGGCTTCTTATGAAACCTTACATCTTATG 354

QY 90 PheLeuPheTyrlleProAlaGlySerTyrllePheTh---GlyTyrlleMetLeu 108
DB 355 TTCTTGGCTTACGCTCCAGCAGCGAGTGGATGACAGTGAAGTGTACTTATG 414

QY 109 GluLeuGlyAlaLeuGlyMetValMetProIleThTyrlleLeuAlaGlySerHis 128
DB 415 AACTTGGAGATTGTCGATTCCTCTTTCGCAATTCGTGTAAAGAGCCACA 474

QY 129 ArgValMetIleValGlyTyrlleCysAlaAlaIleAsnValAlaValPheAlaPro 148
DB 475 CTTGCTAAGATTATGAGAGAAATCTGTCGATTCCTGTTGGTTTTCGCTGCT 534

QY 149 LeuSerIleMetArgGlnValIleTyrlleThTyrlleSerValGluPheMetProPheThLeu 168
DB 535 CTAGCATATACAGACGCTAATAAGACAAAGTGTGATGATGACATGCTTACTTA 594

QY 169 SerLeuPheLeuThLeuCysAlaThMetTrpPhePheTyrlleGlyPhePheTyrlleVal 188
DB 595 TCCTTAACCTTATGATGATGCTGTCATGATGCTTATGATGCTTATGCTTCAAGGAC 654

QY 189 PheTyrlleAlaPheProAsnIleLeuGlyPheLeuPheGlyTyrlleValGlnMetLeu 208
DB 655 ATCTATGCTTCTTCCGATGCTGTTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCT 714

QY 209 TyrlleValTyrlleVal---AspSerTyrlleValleAsp 220
DB 209 TyrlleValTyrlleVal---AspSerTyrlleValleAsp 220

```

Db 715 TACGTTCTCAAACTCTGTAACGTCGCCGATCTAGAGAGAAAGTGAAGCT 774  
Qy 221 Glu1ys-----SerAspPro--- 225  
Db 775 GCTAAGTACCGAGTGAGCCGATATGTTGAAGTACGACAGTTTCATCCCTGAG 834  
Qy 226 -----ValArgGluAlaThrLysSerIysgluGlyValGluIleIleIleAsn 241  
Db 835 CCAATCTCACTGCTTCGTCAAGCAAGTGAACCTGCGGAATGATCGAAGGCTGAG 894  
Qy 242 IleGluAsp 244  
Db 895 ATTGAAGAT 903  
RESULT 11  
AAC41129  
ID AAC41129 standard; DNA; 1189 BP.  
XX  
AC AAC41129;  
XX  
DT 17-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 30755.  
XX  
KM Hybridisation assay; genetic mapping; gene expression control;  
KM protein identification; signal transduction pathway; metabolic pathway;  
KM promoter; termination sequence; ss.  
XX  
OS Arabidopsis thaliana.  
XX  
PN EPI033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-00301439.  
XX  
PR 25-FEB-1999; 99US-0121825P.  
PR 05-MAR-1999; 99US-0123180P.  
PR 09-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 28-APR-1999; 99US-0130891P.  
PR 30-APR-1999; 99US-0131449P.  
PR 04-MAY-1999; 99US-0132407P.  
PR 05-MAY-1999; 99US-0132484P.  
PR 06-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 06-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.

PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 19-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147392P.  
PR 05-AUG-1999; 99US-0147260P.



XX Hybridisation assay; genetic mapping; gene expression control;  
KW protein identification; signal transduction pathway; metabolic pathway;  
KM promoter; termination sequence; ss.  
XX Arabidopsis thaliana.  
XX EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-00301439.  
XX  
XX 25-FEB-1999; 99US-0121825P.  
PR 05-MAR-1999; 99US-0123180P.  
PR 09-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 28-APR-1999; 99US-0130891P.  
PR 30-APR-1999; 99US-0132048P.  
PR 30-APR-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 06-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 18-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138547P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.

PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 30-AUG-1999; 99US-0151303P.

```

PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155559P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161933P.
PR 29-OCT-1999; 99US-0162142P.

```

```

Alignment Scores:
Pred. No.: 3,46e-60 Length: 1219
Score: 621.50 Matches: 110
Percent Similarity: 75.73% Conservative: 46
Best Local Similarity: 53.40% Mismatches: 49
Query Match: 45.93% Indels: 1
DB: 3 Gaps: 1
US-09-743-885a-1 (1-265) x AAC49937 (1-1219)

```

```

QY 9 LeuSerPheIlePheGlyLeuLeuGlyAsnIleValSerPheMetValPheLeuAlaPro 28
DB 190 TTGGCTACGCTCTTTGGCATCTTAGGGAATATCATCTCTTTTGGCTGCTGCCCCA 249
QY 29 ValProThrPheTyrIleValIleTyrIleValIleValIleValIleValIleValIlePro 48
DB 250 ATACCAACGCTCTCTTATATACAGAGAGATCATCTTAAGGATATCATGATCCCT 309
QY 49 TyrMetValAlaIlePheSerAlaGlyLeuLeuLeuTyrTyrAlaTyrLeuAlaGlyAsn 68
DB 310 TACGTGTTCTCTGTTACGGGAGATCTTAGTGTACTACGCAATGATCAAGAAAGAT 369
QY 69 AlaTyrLeuIleValSerIleValIleValIleValIleValIleValIleValIleValSer 88
DB 370 GCTATGATGTGTACTATCACTCATCTCTTGGCTTTGTCTGTAACAATGCTCATCATCTCC 429

```

```

QY 89 LeuPheLeuPheTyrAlaProArgIleValSerIlePheThr---GlyTrpIleMetLeu 107
DB 430 TTATTTCTTTCTTACGCTCTTAAGAGAAAGACTCTAACGGTAGAGTTCCTCTTT 489
QY 108 LeuGluLeuGlyAlaLeuGlyMetValMetProIleThrTyrIleLeuAlaIleGluGlySer 127
DB 490 GTGATGTTTGGGTTGGCCCAATCTTGATGATATCTTATTAATTCATGCCAAT 549
QY 128 HisArgValMetIleValAlaIleTyrIleCysAlaAlaIleAsnValAlaValPheAlaIle 147
DB 550 AACGTTCAAGTCTTGTGATACATTTGATGCTTGTCTTACCTCTTCTTTTGTGT 609
QY 148 ProLeuSerIleMetArgIleValIleValIleValIleValIleValIleValIleValIle 167
DB 610 CCCCTTGGCATTAATGGAAGATGATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 669
QY 168 LeuSerLeuPheLeuThrIleCysAlaIleMetIleTyrPhePheTyrIlePhePheValys 187
DB 670 CTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 729
QY 188 AspPheTyrIleAlaPheProAsnIleLeuGlyPheLeuPheGlyIleValGluMetLeu 207
DB 730 GACATGACATGCCCCGCAATGTTGGTTGATCTTGGAGTACTTACATGATGATT 789
QY 208 LeuTyrPheValTyrIleValys 213
DB 790 CTTTCTTGTATCAAG 807

RESULT 13
AAC39463
ID AAC39463 standard; DNA; 1221 BP.
XX AAC39463;
XX
XX 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 24723.
XX
XX Hybridization assay; genetic mapping; gene expression control;
XX protein identification; signal transduction pathway; metabolic pathway;
XX promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
XX EPI033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-00301439.
XX
XX 25-FEB-1999; 99US-0121825P.
XX 05-MAR-1999; 99US-0123180P.
XX 09-MAR-1999; 99US-0123548P.
XX 23-MAR-1999; 99US-0125788P.
XX 25-MAR-1999; 99US-0126264P.
XX 29-MAR-1999; 99US-0126785P.
XX 01-APR-1999; 99US-0127462P.
XX 06-APR-1999; 99US-0128234P.
XX 08-APR-1999; 99US-0128714P.
XX 16-APR-1999; 99US-0129845P.
XX 19-APR-1999; 99US-0130077P.
XX 23-APR-1999; 99US-0130449P.
XX 23-APR-1999; 99US-0130510P.
XX 23-APR-1999; 99US-0130891P.
XX 28-APR-1999; 99US-0131499P.
XX 30-APR-1999; 99US-0132048P.
XX 30-APR-1999; 99US-0132407P.
XX 04-MAY-1999; 99US-0132484P.
XX 05-MAY-1999; 99US-0132485P.
XX 06-MAY-1999; 99US-0132486P.
XX 06-MAY-1999; 99US-0132487P.
XX 07-MAY-1999; 99US-0132863P.

```

PR 11-MAY-1999;	99US-0134256P.	PR 23-JUL-1999;	99US-0145218P.
PR 14-MAY-1999;	99US-0134218P.	PR 23-JUL-1999;	99US-0145224P.
PR 14-MAY-1999;	99US-0134219P.	PR 26-JUL-1999;	99US-0145276P.
PR 14-MAY-1999;	99US-0134221P.	PR 27-JUL-1999;	99US-0145913P.
PR 14-MAY-1999;	99US-0134370P.	PR 27-JUL-1999;	99US-0145918P.
PR 18-MAY-1999;	99US-0134768P.	PR 27-JUL-1999;	99US-0145919P.
PR 19-MAY-1999;	99US-0134944P.	PR 28-JUL-1999;	99US-0145951P.
PR 20-MAY-1999;	99US-0135124P.	PR 02-AUG-1999;	99US-0146386P.
PR 21-MAY-1999;	99US-0135353P.	PR 02-AUG-1999;	99US-0146388P.
PR 24-MAY-1999;	99US-0136292P.	PR 02-AUG-1999;	99US-0146389P.
PR 25-MAY-1999;	99US-0136021P.	PR 03-AUG-1999;	99US-0147038P.
PR 27-MAY-1999;	99US-0136392P.	PR 04-AUG-1999;	99US-0147204P.
PR 28-MAY-1999;	99US-0136782P.	PR 04-AUG-1999;	99US-0147302P.
PR 01-JUN-1999;	99US-0137222P.	PR 05-AUG-1999;	99US-0147182P.
PR 03-JUN-1999;	99US-0137528P.	PR 05-AUG-1999;	99US-0147260P.
PR 04-JUN-1999;	99US-0137502P.	PR 06-AUG-1999;	99US-0147303P.
PR 07-JUN-1999;	99US-0137724P.	PR 06-AUG-1999;	99US-0147416P.
PR 08-JUN-1999;	99US-0138094P.	PR 09-AUG-1999;	99US-0147493P.
PR 10-JUN-1999;	99US-0138540P.	PR 09-AUG-1999;	99US-0147935P.
PR 10-JUN-1999;	99US-0138847P.	PR 10-AUG-1999;	99US-0148171P.
PR 14-JUN-1999;	99US-0139119P.	PR 11-AUG-1999;	99US-0148319P.
PR 16-JUN-1999;	99US-0139452P.	PR 12-AUG-1999;	99US-0148341P.
PR 16-JUN-1999;	99US-0139453P.	PR 13-AUG-1999;	99US-0148565P.
PR 17-JUN-1999;	99US-0139453P.	PR 13-AUG-1999;	99US-0148684P.
PR 18-JUN-1999;	99US-0139454P.	PR 16-AUG-1999;	99US-0149368P.
PR 18-JUN-1999;	99US-0139455P.	PR 17-AUG-1999;	99US-0149175P.
PR 18-JUN-1999;	99US-0139456P.	PR 18-AUG-1999;	99US-0149465P.
PR 18-JUN-1999;	99US-0139457P.	PR 20-AUG-1999;	99US-0149722P.
PR 18-JUN-1999;	99US-0139458P.	PR 20-AUG-1999;	99US-0149723P.
PR 18-JUN-1999;	99US-0139459P.	PR 20-AUG-1999;	99US-0149929P.
PR 18-JUN-1999;	99US-0139460P.	PR 23-AUG-1999;	99US-0149992P.
PR 18-JUN-1999;	99US-0139461P.	PR 23-AUG-1999;	99US-0149930P.
PR 18-JUN-1999;	99US-0139462P.	PR 25-AUG-1999;	99US-0150566P.
PR 18-JUN-1999;	99US-0139463P.	PR 26-AUG-1999;	99US-0150884P.
PR 18-JUN-1999;	99US-0139750P.	PR 27-AUG-1999;	99US-0151065P.
PR 18-JUN-1999;	99US-0139763P.	PR 27-AUG-1999;	99US-0151066P.
PR 21-JUN-1999;	99US-0139817P.	PR 27-AUG-1999;	99US-0151080P.
PR 22-JUN-1999;	99US-0139899P.	PR 30-AUG-1999;	99US-0151303P.
PR 23-JUN-1999;	99US-0140353P.	PR 31-AUG-1999;	99US-0151438P.
PR 24-JUN-1999;	99US-0140354P.	PR 01-SEP-1999;	99US-0151930P.
PR 24-JUN-1999;	99US-0140695P.	PR 07-SEP-1999;	99US-0152363P.
PR 28-JUN-1999;	99US-0140823P.	PR 10-SEP-1999;	99US-0153070P.
PR 29-JUN-1999;	99US-0140991P.	PR 13-SEP-1999;	99US-0153758P.
PR 30-JUN-1999;	99US-0141287P.	PR 15-SEP-1999;	99US-0154018P.
PR 01-JUL-1999;	99US-0141842P.	PR 16-SEP-1999;	99US-0154039P.
PR 01-JUL-1999;	99US-0142154P.	PR 20-SEP-1999;	99US-0154779P.
PR 02-JUL-1999;	99US-0142055P.	PR 22-SEP-1999;	99US-0155139P.
PR 06-JUL-1999;	99US-0142390P.	PR 23-SEP-1999;	99US-0155486P.
PR 08-JUL-1999;	99US-0142803P.	PR 24-SEP-1999;	99US-0155659P.
PR 09-JUL-1999;	99US-0142920P.	PR 28-SEP-1999;	99US-0156488P.
PR 12-JUL-1999;	99US-0142977P.	PR 29-SEP-1999;	99US-0156596P.
PR 13-JUL-1999;	99US-0143542P.	PR 04-OCT-1999;	99US-0157117P.
PR 14-JUL-1999;	99US-0143624P.	PR 05-OCT-1999;	99US-0157753P.
PR 15-JUL-1999;	99US-0144005P.	PR 06-OCT-1999;	99US-0157665P.
PR 16-JUL-1999;	99US-0144085P.	PR 07-OCT-1999;	99US-0158029P.
PR 16-JUL-1999;	99US-0144086P.	PR 08-OCT-1999;	99US-0158232P.
PR 19-JUL-1999;	99US-0144325P.	PR 12-OCT-1999;	99US-0158369P.
PR 19-JUL-1999;	99US-0144331P.	PR 13-OCT-1999;	99US-0159293P.
PR 19-JUL-1999;	99US-0144332P.	PR 13-OCT-1999;	99US-0159294P.
PR 19-JUL-1999;	99US-0144333P.	PR 13-OCT-1999;	99US-0159295P.
PR 19-JUL-1999;	99US-0144334P.	PR 14-OCT-1999;	99US-0159329P.
PR 19-JUL-1999;	99US-0144335P.	PR 14-OCT-1999;	99US-0159330P.
PR 20-JUL-1999;	99US-0144352P.	PR 14-OCT-1999;	99US-0159331P.
PR 20-JUL-1999;	99US-0144884P.	PR 14-OCT-1999;	99US-0159637P.
PR 21-JUL-1999;	99US-0144814P.	PR 14-OCT-1999;	99US-0159638P.
PR 21-JUL-1999;	99US-0145086P.	PR 18-OCT-1999;	99US-0159584P.
PR 21-JUL-1999;	99US-0150888P.	PR 21-OCT-1999;	99US-0160741P.
PR 22-JUL-1999;	99US-0145085P.	PR 21-OCT-1999;	99US-0160767P.
PR 22-JUL-1999;	99US-0145087P.	PR 21-OCT-1999;	99US-0160768P.
PR 22-JUL-1999;	99US-0145089P.	PR 21-OCT-1999;	99US-0160770P.
PR 22-JUL-1999;	99US-0145192P.	PR 21-OCT-1999;	99US-0160815P.
PR 23-JUL-1999;	99US-0145145P.	PR 22-OCT-1999;	99US-0160980P.







```
QY 70 TyrLeuIleValSerIleAsnGlyPheGlyCysAlaIleGluLeuThrTyrIleSerLeu 89
Db 338 TTCCTTCTGTGTCACCAACACCTTGGTGGCTTCAACCAATATCATCTCCATC 397
QY 90 PheLeuPheTyrAlaProArgIlySerIlePheThr--GlyTyrLeuMetLeuLeu 108
Db 398 TTGTGTGCTCCATCCATCAAGAAAGCCGATCTAAGCGTGAAGCTTGTGTAAG 457
QY 109 GluLeuGlyAlaLeuGlyMetValMetProIleThrTyrIleLeuGluGlySerHis 128
Db 458 AACTTGGAGGGTTCGTGGTATCTCCCTCCATTCCTGCAAAAGGAACACCA 517
QY 129 ArgValMetIleValGlyTyrPheCysAlaAlaIleAsnValAlaValPheAlaPro 148
Db 518 CGGCGCAAGATCATTTGAGGATCTGTGTGCGATCTCTGTGCGTTTGTGCGCCG 577
QY 149 LeuSerIleMetArgGlnValIleIlySerThrIlySerValGluPheMetProPheThrLeu 168
Db 578 CTAGCATATATCAGAACGGGATAAAGACGAAAGTGTGAGTACATGCGCTTACCTTA 637
QY 169 SerLeuPheLeuThrLeuCysAlaThrMetTyrPhePheTyrGlyPhePheIlySerAsp 188
Db 638 TCCCTGACCTTACCATCAGTGGGTCATATGGCTCCTTATGCTCTTAAAGAT 697
QY 189 PheTyrIleAlaPheProAsnIleLeuGlyPheLeuPheGlyIleValGlnMetLeuLeu 208
Db 698 ATCTATGTCCTCCCTCCCAACGATGATGGGTTGTTCTAGTGACACTTCAATGATATC 757
QY 209 TyrPheValTyrIlyS-----AspSer 215
Db 758 TATGTGTTTACCAATATGCAAAACGCCGTCGATTTGGTGAAGAACTTGAGGCT 817
QY 216 LysArgIleAspAspGluIlySerAspProValArgGluAlaThr--LysSerIlyGlu 234
Db 818 GCGAAATTCGCAGAGTGAAGCATCATATGATGAGTGAAGTACCTCATCTCTGAA 877
QY 235 GlyValGluIleIleIle-----AsnIleGluAspAspAsnSer 247
Db 878 CCGATGCGATCACCGTCGTCGATCGGTGACGACATCATCTAAGCATCAAAATGCT 937
QY 248 Asp 248
Db 938 GAG 940

RESULT 15
AAC47147
ID AAC47147 standard; DNA; 1252 BP.
XX AAC47147;
AC
XX
XX 18-OCT-2000 (first entry)
DT
XX
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 52735.
DE
XX
XX Hybridisation assay; genetic mapping; gene expression control;
KM protein identification; signal transduction pathway; metabolic pathway;
KM promoter; termination sequence; ss.
XX Arabidopsis thaliana.
OS
XX
XX EP1033405-A2.
PN
XX
XX 06-SEP-2000.
PD
XX
XX 25-FEB-2000; 2000EP-00301439.
PF
XX
XX 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
```

```
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0131449P.
PR 04-MAY-1999; 99US-0132048P.
PR 05-MAY-1999; 99US-0132484P.
PR 06-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 07-MAY-1999; 99US-0132487P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 18-MAY-1999; 99US-0134370P.
PR 19-MAY-1999; 99US-0134768P.
PR 20-MAY-1999; 99US-0134941P.
PR 21-MAY-1999; 99US-0135124P.
PR 24-MAY-1999; 99US-0135353P.
PR 25-MAY-1999; 99US-0135629P.
PR 27-MAY-1999; 99US-0136021P.
PR 28-MAY-1999; 99US-0136392P.
PR 01-JUN-1999; 99US-0136782P.
PR 03-JUN-1999; 99US-0137222P.
PR 04-JUN-1999; 99US-0137528P.
PR 07-JUN-1999; 99US-0137502P.
PR 08-JUN-1999; 99US-0137724P.
PR 10-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 14-JUN-1999; 99US-0138847P.
PR 16-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 17-JUN-1999; 99US-0139453P.
PR 18-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140911P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
```

PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144332P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145152P.  
PR 23-JUL-1999; 99US-0145152P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 28-JUL-1999; 99US-0145918P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 05-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147302P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 16-AUG-1999; 99US-0148684P.  
PR 17-AUG-1999; 99US-0149368P.  
PR 18-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 30-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.

PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161922P.  
PR 28-OCT-1999; 99US-0161933P.  
PR 29-OCT-1999; 99US-0162142P.

## Alignment Scores:

Pred. No.:	1,31e-56	Length:	1252
Score:	590.00	Matches:	126
Percent Similarity:	66.41%	Conservative:	46
Best Local Similarity:	48.65%	Mismatches:	57
Query Match:	43.61%	Indels:	30
DB:	3	Gaps:	6

US-09-743-885a-1 (1-265) x AAC7147 (1-1252)

QY 10 SerPheIlePheGlyLeuLeuGlyAenIleValSerPheMetValPheLeuAlaProVal 29  
Db 189 GCATTGTGTTGGATTGCGGTAACTATCATATCATGCTCGTGGTCCGCGCCAGTG 248  
QY 30 ProThPheTyriyIleTyriyAArglySerSerGluGlyTyriGlnAlaIleProTy 49  
Db 249 CCCACTTTGTAAAGATATCAAGAAATCAACGAAAGTTTCAGTCTTACCTAT 308  
QY 50 MetValAlaLeuPheSerIleGlyLeuLeuTyriyAlaTyriyLeuAArgly--Asn 68  
Db 309 GTTTCAGCACTTTTTCGCGGAGCTTTGGATTACTACCTATGCAAAAGATGCACA 368  
QY 69 AlaTyriLeuIleValSerIleAenglyPheGlyCysAlaIleGluLeuThyTyriIleSer 88  
Db 369 GCCTTCTTCTATCAACCATTAACGCTTTGGATGGTATCAAAACCATCTACATGTC 428  
QY 89 LeuPheLeuPheTyriyAlaProAArglySerIlePheThr-----GlyTyr 104  
Db 429 CTCTTGTCTCCATATGCTAAAGAAATCAAGATATCCACTTGAAGTTCTTGCTTC 488  
QY 105 LeuMetLeuLeuIleuGlyAlaLeuGlyMetValMetProIleThrTyriLeuAla 124  
Db 489 TTGAACCTTTTGGGATTTGCCGCAATGCTTCTGTCGAG-----CTCTTAACC 539  
QY 125 GluGlySerIleAArgValMetIleValGlyTyrIleCysAlaIleAenValAlaVal 144  
Db 540 AAAGTTTCAACAGTGAAGAAAGTTTCGAGGAGATTTCGTTGGATTTCGCAAGTT 599  
QY 145 PheAlaIleProLeuSerIleMetAArgIleValIleTyriySerValGluPheMet 164  
Db 600 TTGCGAGCTCTTGTGATATCATGAGAGTGTGTGCAACAAAGATGTGAGATTATG 659  
QY 165 ProPheThriLeuSerIleuPheLeuThriLeuCysAlaIleThMetThrPhePheTyriGlyPhe 184  
Db 660 CTTTCTTATATCGTGTGTTTCTTACATTAGCGCGCTCATGTCGCTCTTACGGGTCTC 719

```
Qy 185 PheLyLeuSerPheTyrIleAlaPheProAsnIleuGlyPheLeuPheGlyIleVal 204
    |||||
Db 720 GCTATTAAAGACTTCTACGTTGCGCTTCCAAATGATTTAGCGCGCTTCTTAGAGACTGTT 779
    |||||
Qy 205 GlMetLeuLeuTyrPheValTyrLyAspSerLySArgIleAspAspGlyLySerAsp 224
    |||||
Db 780 CAATGATCTCTACATAATTTCAAAATACTACAAAACCTCGTGGCTCAGAAACAGAT 839
    |||||
Qy 225 ProValArgGluAlaThrLySerLySArgGlyValGluIleIleIleAsnIleGluAsp 244
    |||||
Db 840 -----AAATCCAA----- 848
    |||||
Qy 245 AspAsnSerAspAsnAlaLeuGlnSerMetGlyLyAspPheSerArgLeuArgThr 263
    |||||
Db 849 GACGTGTCGATCATTCATC-----GACATAGCAAAAGCTAACGACC 890
    |||||
```

Search completed: December 24, 2004, 21:04:22  
Job time : 596 secs

**this Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: December 24, 2004, 21:04:27 : Search time 5448 Seconds  
(without alignments)  
2300.252 Million cell updates/sec

Title: US-09-743-885A-1  
Perfect score: 1353  
Sequence: 1 MAQLRADDSTFGLGNTV.....NSDNALQSMKEDSRRTSK 265

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 4526729 seqs, 23644849745 residues  
Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0  
Maximum DB seq length: 200000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODE=frame+ p2n.model -DEV=xlp  
-Q=/cgn2.1/USPTO.spool.p/US09743885/runat.23122004.165258.9632/app.query.fasta.1.455  
-DB=GenEmbl-QFMT=faaap -SUFFIX=gcg -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdd -LIST=45  
-DOCAIGN=200 -THR.SCOR=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=pcr -NOR=exc -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09743885\_QCGN.1.1.7406 @runat.23122004.165258.9632 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG.SCOR=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl.\*  
1: gb\_ba:\*  
2: gb\_hcg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pac:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vi:\*  
  
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	1353	100.0	1182	8	AF313914 Petunia x
2	1353	100.0	1205	6	BD217625 Method of
3	1353	100.0	1205	6	AX006355 Sequence
4	703.5	52.0	902	8	AY086047 Arabidops

5	656	48.5	1117	8	MTN3	Y08726 M.truncatui
6	652	48.2	1428	8	BT013320	BT013320 Lycopersi
7	647	47.8	1562	8	AK103266	AK103266 Oryza sat
8	646.5	47.8	1494	8	AK101913	AK101913 Oryza sat
9	623.5	46.1	879	6	AX506568	AX506568 Sequence
10	623.5	46.1	910	8	AY113934	AY113934 Arabidops
11	623.5	46.1	1269	8	AY087836	AY087836 Arabidops
12	623.5	46.1	1291	8	AY045949	AY045949 Arabidops
13	622	46.0	670	6	CO805384	CO805384 Sequence
14	622	46.0	870	8	AY078041	AY078041 Arabidops
15	622	46.0	901	8	AY096594	AY096594 Arabidops
16	622	46.0	1157	8	AF361825	AF361825 Arabidops
17	622	46.0	1172	8	AF419559	AF419559 Arabidops
18	622	46.0	1176	8	BT000808	BT000808 Arabidops
19	622	46.0	1196	8	AY070412	AY070412 Arabidops
20	621.5	45.9	934	8	AY064674	AY064674 Arabidops
21	621.5	45.9	1221	8	AY084672	AY084672 Arabidops
22	621.5	45.9	1250	8	AY054548	AY054548 Arabidops
23	618	45.7	889	8	AY059108	AY059108 Arabidops
24	618	45.7	1124	8	AY116672	AY116672 Arabidops
25	618	45.7	1209	8	AF095641	AF095641 Arabidops
26	618	45.7	1279	8	AY057575	AY057575 Arabidops
27	590	43.6	1252	8	AY087516	AY087516 Arabidops
28	571.5	42.2	1492	8	AK109114	AK109114 Oryza sat
29	539	39.8	1500	8	AK106127	AK106127 Oryza sat
30	535	39.5	1517	8	AK070510	AK070510 Oryza sat
31	535	39.5	128673	8	AP005299	AP005299 Oryza sat
32	523.5	38.7	96827	8	AC005770	AC005770 Arabidops
33	484	35.8	134938	2	AC137759	AC137759 Oryza sat
34	474.5	35.1	82207	8	AT119155	AT119155 Arabidops
35	467.5	34.6	66087	8	AB023037	AB023037 Arabidops
36	463.5	34.3	2667	8	MT8307887	MT8307887 Medicago
37	462.5	34.2	154826	8	CNS079P4	AL713927 Oryza sat
38	462	34.1	944	8	BT002983	BT002983 Arabidops
39	447	33.0	775	8	BT005476	BT005476 Arabidops
40	447	33.0	960	8	BT004185	BT004185 Arabidops
41	447	33.0	990	8	AY088391	AY088391 Arabidops
42	440.5	32.6	1161	8	AK099531	AK099531 Oryza sat
43	440.5	32.6	1208	8	AK067960	AK067960 Oryza sat
44	435.5	32.2	693	8	BT100528	BT100528 Arabidops
45	425	31.4	1211	8	AK063475	AK063475 Oryza sat

## ALIGNMENTS

RESULT 1  
AF313914 1182 bp mRNA linear PLN 05-JUL-2002  
LOCUS Petunia x hybrida NECL mRNA, complete cds.  
DEFINITION AF313914  
ACCESSION AF313914  
VERSION AF313914.1 GI:11345412  
KEYWORDS  
SOURCE  
ORGANISM  
Petunia x hybrida  
Petunia x hybrida  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
asterids; Jamidae; Solanales; Solanaceae; Petunia.  
REFERENCE  
1 (bases 1 to 1182)  
Ge,Y.-X., Angenent,G.C., Wittlich,P.E., Peter,J., Franken,J., and  
Buescher,M., Zhang,L.-M., Dahlhaus,E., Kater,M.M., Willems,G.J. and  
Creemers-Molenaar,T.  
NECL, a novel gene, highly expressed in nectary tissue of Petunia  
hybrida  
JOURNAL Plant J. 24 (6), 725-734 (2000)  
MEDLINE 20575716  
PUBMED 11135107  
REFERENCE  
2 (bases 1 to 1182)  
Ge,Y.-X., Angenent,G.C., Dahlhaus,E., Franken,J., Willems,G.J. and  
Creemers-Molenaar,T.  
Partial gene silencing of NECL results in early opening of anthers  
in Petunia hybrida  
JOURNAL Mol. Gen. Genet. (2001) In press  
REFERENCE  
3 (bases 1 to 1182)

AUTHORS Creemers-Molenaar,T., Ge,Y.-X. and Angenent,G.C.  
 TITLE Direct Submission  
 JOURNAL Submitted (17-OCT-2000) Plant Development and Reproduction, Plant Research International, Droevendaalsesteeg 1, Wageningen 6708 PB, The Netherlands

## FEATURES

## source

Location/Qualifiers  
 1..1182

## CDS

/organism="Petunia x hybrida"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:4102"  
 56..853  
 /note="nectary-specific"  
 /codon\_start=1  
 /product="NEC1"  
 /protein\_id="AAC34696.1"  
 /db\_xref="GI:11345413"  
 /translation="MAQLRADDLISFIFGLGNIVSFNVFLAPVPTFYKIKKSSRGYQAIRYMALBSAGLLIYYAIRKRAYIIVSINGGCAIELTYSIFLPYAPRSKIRTGMLMLBGLAGMVPITYLIAESGSHVMIVGMCAINAVAPUSIRQVTKTSSVEFMPTLSFLTLCATMWFYGFKKDPIAPNIIIGFLGVOMLIFYVYKDSKRI DDEKSDVREATKSKSEVEIINIENDDNSDALQSMKDFSLRTSK"

## ORIGIN

## Alignment Scores:

Pred. No.: 5.95e-116 Length: 1182  
 Score: 1353.00 Matches: 265  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 8 Gaps: 0

US-09-743-885A-1 (1-265) x AF313914 (1-1182)

Qy 1 MetlaaglnleuargAlaaspAspleuSerPheIllePheglyleuLeuglYAsnIlleVal 20  
 Db 56 ATGGCGCAATTAACGTGCTGATGACTGCTTTCATTTATTTGGCCCTTCTGGTAATATTGTA 115  
 Qy 21 SerPheMetValPheleuAlaProValProTherPheTyrlYsIlleTyrlYsArgLysSer 40  
 Db 116 TCATTCATGCTGCTTCCTAGACCCGTCGCAACATTTTACAAATATATATAAAGAAATCA 175  
 Qy 41 SerGluglYTyrglnAlaIlleProTyMetValAlaleuPheSerAlaaglYleuLeu 60  
 Db 176 TCAGAGAGATATCAAGCAATATCATATATGGTACACATTCACGCGGACCTATTGCTA 235  
 Qy 61 TyrtYrAlaTyrlleuArgLysAsnAlaTyrlleuIlleValSerIlleAsnGlYPhaGlYCyS 80  
 Db 236 TATTATGCTTATCTCAGGAAGATGCCATCTTATCGACGATTAATGGCTTTGGATGT 295  
 Qy 81 AlaIllegluleuThrTyrlleSerleuPheleuPheTyrlAlaProArgLysSerLysIlle 100  
 Db 296 GCCATTGAATTAACATATATCTCTGTTTCTCTTTTACGCGCCAGAAAGCTTAAGATT 355  
 Qy 101 PheThrglYTrpleuMetleuLeuGluleuGlYAlaleuGlYMetValMeProlleThr 120  
 Db 356 TTTCAGAGGTGGCTGATGCTCTTAGAATGGAGCCCTCAGAAAGTGATGCCAATTACT 415  
 Qy 121 TyrlleuAlaaglYSerHlsArgValMetIlleValaglYTrpleCyAlaAlaIlle 140  
 Db 416 TATTATTAGCAAGAGCTCATAGACTGATGATGGGATGGATTGGCGACTATC 475  
 Qy 141 AsnValAlaValPheAlaAlaProlleuSerIlleMetArglnValIlleYsThrLysSer 160  
 Db 476 AATGTTGCTGCTTTGGCTGCTCTTTAAGCATCATGAGGCAAGTAATTAACAAAGAGT 535  
 Qy 161 ValGlupheMetProPheThrlleuSerleuPheleuThrlleuCyAlaIthMetTrpPhe 180  
 Db 536 GTAAGTTCATGCCCTTCACTTATCTTGTCTCATCTCTGTGCCACTATATGGGTTT 595  
 Qy 181 PheTyrglYPhePheLysLysAspPheTyrlleAlaPheProAsnIlleLeuGlYPheleu 200  
 Db 596 TTCTATGGGTTTTCAGAAAGACTTTTACATTCGCTTTCCAAATATATCTGGGCTTTCTA 655

Qy 201 PheglYlIleValgluMetleuLeuTyrlPheValTyrlYsAspSerLysArgIlleAspAsp 220  
 Db 656 TTGGGAATCGTTCAAAATGCTATATATATTTGTTTACAAAGATTCAAGAAATATAGATAT 715  
 Qy 221 GlulYSerAspProValArgGlAlaThrlYsSerleuGluglYValglulleIlle 240  
 Db 716 GAAATAATCATCTGTTGCAAGAGCTACAAATCAAAAGAGGTGAGAAATCATATATC 775  
 Qy 241 AsnIlleGlulAspAsnSerAspAsnAlaLeuGlInSerMetGlulYAspPheSerArg 260  
 Db 776 AACATTGAATGATATATTCTGATTAACGATTCACATTCAGTGAAGAAAGATTTTCCAGA 835  
 Qy 261 LeuArgThrSerLys 265  
 Db 836 CTCGGGACATCAAAA 850

## RESULT 2

## LOCUS

BD217625 1205 bp DNA linear PAT 17-JUL-2003  
 Method of collecting metabolite from modified nectar by insect.

## DEFINITION

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

Petunia x hybrida  
 Petunia x hybrida  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 asterids; lamiales; Solanales; Solanaceae; Petunia.

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

Creemers,J., Angenent,G.C. and Kater,M.M.  
 Method of collecting metabolite from modified nectar by insect  
 Patent: JP 2002520064-A 1 09-JUL-2002;  
 STITCHING CENTRUM VOOR PLANTENVERBEDLINGEN EN REPRODUKTIEONDERZOEK

## COMMENT

## OS

## PN

## PD

## PF

## PR

## PJ

## PC

## CC

## CC

## CC

## CC

## CC

## CC

## CC

## CC

## CC

## CC

## CC

## CC

## CC

## CC

## CC

## CC

## CC

## CC

## CC

## CC

## CC

## CC

## CC

## CC

## CC

## CC

## CC

## CC

## CC

## CC

## CC

## CC

## CC

## CC

## CC

## CC

## CC

## CC

## CC

## CC

## CC

## CC

## ORIGIN

## Alignment Scores:

Pred. No.: 6.08e-116 Length: 1205  
 Score: 1353.00 Matches: 265  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 6 Gaps: 0

US-09-743-885A-1 (1-265) x BD217625 (1-1205)

Qy 1 MetlaaglnleuargAlaaspAspleuSerPheIllePheglyleuLeuglYAsnIlleVal 20  
 Db 79 ATGGCGCAATTAACGTGCTGATGACTGCTTTCATTTATTTGGCCCTTCTGGTAATATTGTA 138  
 Qy 21 SerPheMetValPheleuAlaProValProTherPheTyrlYsIlleTyrlYsArgLysSer 40  
 Db 139 TCATTCATGCTCTTCCTAGACCCGTCGCAACATTTTACAAATATATATAAAGAAATCA 198  
 Qy 41 SerGluglYTyrglnAlaIlleProTyMetValAlaleuPheSerAlaaglYleuLeu 60

## FEATURES

## source

Location/Qualifiers  
 1..1205  
 /organism="Petunia x hybrida"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:4102"

[illegible]

		/notes="unnamed protein product; NECL"	
		/codon_start=1	
		/protein_id="CAC07355.1"	
		/db_xref="GI:9994502"	
		/translation="MAQLRADDLSPFILLGNIVSFVFLAPVPTFYKIYKRKSEGYQAIYVYVALFSAGLLLYVAYLRKNAVIVSINGSCAIILELYISLFLPYARSKRIFGMILVELGALGMVPIITVYLAIEGSHVMIIGMCALINAVFAPAPISIMPOVKTKEVEMPPPLSLFLTICATMMPEYFGFKPDYAPFNIIIGPLGIOMLIPFYKOSKRIRDDEKSDPVRERATRSKEGEVLIINIEDNSDNALOSMEKDSRLATSK"	
ORIGIN			
Alignment Scores:			
Pred. No.:	6,08e-116	Length:	1205
Score:	1353.00	Matches:	265
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0
US-09-743-885A-1 (1-265) x AX006355 (1-1205)			
QY	1	MetAaGInLeuAaGAlaAspAaPLeuSerPheIlePheGlyLeuLeuGlyAsnIleVal	20
Db	79	ATGGGGCAATTAACGCTGATGACCTTGCTTCAATATTGGCCCTTCTGGTAATATGTA	138
QY	21	SerPheMetValPheLeuAlaProValProThrPheTyrIleTyrIleTyrAspIleSer	40
Db	139	TCATTCAAGGCTCTTCACGACCCGGTGCACATTTTACAAATAATATTAAGGAATCA	198
QY	41	SerGInGlyTyrGInAlaIleProTyrMetValAlaLeuPheSerAlaGlyLeuLeuLeu	60
Db	199	TCAGAGGATATCAAGCAATTCATATATGTTACACACTGTTCAAGCGCGCACTATTGCTA	258
QY	61	TyrTyrAlaTyrLeuAaGlyAsnAlaTyrLeuIleValSerIleAsnGlyPheGlyCys	80
Db	259	TATTATGCTTATTCACAGAAAGAGCCATCTTATTCGCACACATTAATGCGCTTGGAGT	318
QY	81	AlaIleGInLeuThrTyrIleSerLeuPheLeuPheTyrAlaProAspGlySerIleIle	100
Db	319	GCCATTGAATTAACATATATCTCTCTGTTTCTTTTACGGGCCAGAAAGCTTAAGATT	378
QY	101	PheThrGlyTyrPLeuMetLeuLeuGluLeuGlyAlaLeuGlyMetValMetProIleThr	120
Db	379	TTCAAGAGGTGGCTGATGCTCTTGAATTTGGAGCCCTAGGAATGGTGATGCCAATTACT	438
QY	121	TyrLeuLeuAlaGInGlySerHisAryValMetIleValGlyTyrPileCysAlaAlaIle	140
Db	439	TATTATTATTAAGACAGAGCTCACATAGATGATGATGGATGATTTTGACAGTATC	498
QY	141	AsnValAlaValPheAlaAlaProLeuSerIleMetArgGInValIleTyrThrIleSer	160
Db	499	AATGTGCTGCTCTTGCTGCTCTTTAAGCATCATGAGCGAAGTAAATTAACAAGAGT	558
QY	161	ValGInPheMetProPheThrLeuSerLeuPheLeuThrLeuCysAlaIleMetTyrPhe	180
Db	559	GTAGAGTTCATGCCCTTCACCTTATCTTGTTCCTCACTCTGTGCACTATGTGGATT	618
QY	181	PheTyrGlyPhePheIleValAspPheTyrIleAlaPheProAsnIleLeuGlyPheLeu	200
Db	619	TTCTAATGGGTTTTCAAGAAAGACTTTTACATTCGCTTCCAAATATATCGGGCTTCTTA	678
QY	201	PheGlyIleValGInMetLeuLeuTyrPheValTyrIleAspSerIleAspArgIleAspAsp	220
Db	679	TTCCGAATCGTTCAAATGCTATTTATTTTCTTTACAAAGATTCAAAGAGATGATGAT	738
QY	221	GluIleSerAspProValArgGluAlaThrIleSerIleGlyIleValGluIleIleIle	240
Db	739	GAAAATTCGATCTCTGTTCCAGAAAGCTACAAATAAAGAAAGGTGTAAGAAATCATATC	798
QY	241	AsnIleGluAspAspAsnSerAspAsnAlaLeuGInSerMetGluIleAspPheSerArg	260
Db	799	AACATGTGAAGTAAATCTGATTAAGCATGACATGACATGAGAGAGGATTTTCCAGA	858





TITLE Use of a subtractive hybridization approach to identify new  
 JOURNAL Medicago truncatula genes induced during root nodule development  
 MEDLINE Mol. Plant Microbe Interact. 9 (4), 233-242 (1996)  
 PUBMED 96212994  
 REFERENCE 8634476  
 2 (bases 1 to 1117)  
 AUTHORS Gamas, P.  
 TITLE Direct Submission  
 JOURNAL Submitted (08-OCT-1996) P. Gamas, CNRS-INRA, LBRPM, Chemin De  
 Borda Rouge BP27, Castanet-Tolosan Cedex, 31326, FRANCE  
 FEATURES  
 source Location/Qualifiers  
 1..1117  
 /organism="Medicago truncatula"  
 /mol\_type="mRNA"  
 /cultivar="Jemalong"  
 /db\_xref="taxon:3880"  
 /feature\_type="root node"  
 /note="genotype J5"  
 1..1117  
 /gene="MCN3"  
 115..921  
 /gene="MCN3"  
 /codon\_start=1  
 /protein\_id="CA69976.1"  
 /db\_xref="GI:1619602"  
 /db\_xref="GOA:P93332"  
 /db\_xref="TrEMBL:P93332"  
 /translation="MAISHNTLAFTRGMIGNVISELFLAPISPIFYRIKKSTESFQ  
 SLPLVALPSSMLLYALKKDAFLITNSFGCVETIYIILYIAPRDAIRKTS  
 KLSAMNVGSFALLIVTYAVHGRVGVCSLSVPAELSLVAOVRNRTS  
 VEFPMPLSTLTLSATMWFYGFPLKDCIILPNLGYLGLMLLVAIYANGSEK  
 AMKKEKAPLEPKSIVIEIQEKIEQKKNDDBEKDKSEPIGCGV"  
 115..204  
 /gene="MCN3"  
 /note="putative"  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 1 69e-51 Length: 1117  
 Score: 656.00 Matches: 129  
 Percent Similarity: 68.80% Conservative: 54  
 Best Local Similarity: 48.50% Mismatches: 55  
 Query Match: 48.48% Indels: 28  
 Gaps: 6  
 US-09-743-885a-1 (1-265) x MTN3 (1-1117)  
 QY 9 LeuSerPheIlePheGlyLeuLeuGlyAsnIleValSerPheMetValPheLeuAlapro 28  
 DB 136 TTGGCTTTACCTTGGCATGCTTGGTAACGTCATTTCGTTCTTGGTATTCCTGGCTCCA 195  
 QY 29 ValProThrPheTyrIleValSerPheGlySerSerGluGlyTyrGlnAlaIlePro 48  
 DB 196 ATATCAACATTTTACCGGTATATACAGAAAGAAATCAACAGAGGTTTTCAGTCACACCT 255  
 QY 49 TyrMetValAlaLeuPheSerAlaGlyLeuLeuTyrTyrAlaTyrLeuArgIleAsn 68  
 DB 256 TACCTAGTAGCTTGTTCAGCTTCATGCTTGGTGTGACATCATTCATTCCTCAAAAAGAT 315  
 QY 69 AlaTyrLeuIleValSerIleAsnGlyPheGlyCysAlaIleGluLeuThrTyrIleSer 88  
 DB 316 GCTTTCTCCTCATTCATCAATCAATCATTCATTCATTCATTCATTCATTCATTCATTCAT 375  
 QY 89 LeuPheLeuPheTyrAlaProArgIleSerValIlePheThrGlyTyrLeuMet---Leu 107  
 DB 376 TTGTATCATCATCATCATCAACAGAGATGCAAGAACTTAATTCCTCAAGTTACTTCGCGCA 435  
 QY 108 LeuGluLeuGlyAlaLeuGlyMetValMetProIleThrTyrIleLeuAlaGluGlySer 127  
 DB 436 ATGAATGTGGGGCTTGTGCTTGTATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 495  
 QY 128 HlsArgValMetIleValGlyTyrIleCysAlaAlaIleAsnValAlaValPheAlaAla 147

DB 496 CTTCGTGTCCAAAGTTCTTGATGAGTGTGTGTTCCCTTAGTGTAGTGTGTTGACGACA 555  
 QY 148 ProlSerIleIleMechArgIleValIleIleValIleValIleValIleValIleValIleVal 167  
 DB 556 CCATCAAGCATTTGCTCAAGTGTGCTCAAGTGTGCTCAAGTGTGCTCAAGTGTGCTCAAGTGT 615  
 QY 168 LeuSerLeuPheLeuThrLeuCysAlaThrMetTrpPhePheTyrGlyPhePheValIle 187  
 DB 616 TTGTATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 675  
 QY 188 AspPheTyrIleAlaPheProAlaIleLeuGlyPheLeuPheGlyIleValGlnMetLeu 207  
 DB 676 GACATATGATCATGTTTACCAATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 735  
 QY 208 LeuTyrPheValTyrIleAsp-----SerIleValIleAspGluIleSerAspPro 225  
 DB 736 CTGTATGCAATTTTACAGAAATGCTGGGGAAGAAAGCTATATAAGAAAGAAAGAACACCA 795  
 QY 226 ValArgGluAlaThrIleSerIleGluGlyValGluIleIleIle----- 240  
 DB 796 ATT-----GAGCCACCCAAAGT-----ATTGTTATAGAGACCCCAATTGAG 837  
 QY 241 -----AsnIleGluAspAspAsnSerAspAsnAlaLeuGlnSer 253  
 DB 838 AAAATTGAACAAGAAAGAAAGAAAGAAAGATGATGATGATGATGATGATGATGATGATGAT 879  
 QY 254 MetGluIleValAspPheSer 259  
 DB 880 ---GAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 894  
 RESULT 6  
 BT013320 1428 bp mRNA linear PLN 11-MAY-2004  
 LOCUS Lycopersicon esculentum clone J35010P, mRNA sequence.  
 DEFINITION BT013320  
 VERSION BT013320.1 GI:47104735  
 KEYWORDS FLI CDNA.  
 SOURCE Lycopersicon esculentum (tomato)  
 ORGANISM Lycopersicon esculentum  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 asterids; lamiales; Solanales; Solanaceae; Solanum; Lycopersicon.  
 1 (bases 1 to 1428)  
 REFERENCES  
 Kirkness, E.F., Wang, W. and Vazelle, A.  
 Direct Submission  
 Submitted (11-MAY-2004) The Institute for Genomic Research, 9712  
 Medical Center Drive, Rockville, MD 20850, USA  
 FEATURES  
 source Location/Qualifiers  
 1..1428  
 /organism="Lycopersicon esculentum"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:4081"  
 /clone="J35010P"  
 /issue\_type="mixed elicitor"  
 /note="TMEAM28"  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 5.29e-51 Length: 1428  
 Score: 652.00 Matches: 128  
 Percent Similarity: 70.40% Conservative: 48  
 Best Local Similarity: 51.20% Mismatches: 66  
 Query Match: 48.19% Indels: 8  
 Gaps: 3  
 US-09-743-885a-1 (1-265) x BT013320 (1-1428)  
 QY 10 SerPheIlePheGlyLeuLeuGlyAsnIleValSerPheMetValPheLeuAlaproVal 29  
 DB 193 GCTTTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCT 252  
 QY 30 ProThrPheTyrIleValIleTyrIleValSerSerGluGlyTyrGlnAlaIleProTyr 49

```

Db      253 CCCACGTTCTATGACATTACAGAGAAATCAAGAGGCTATCAATCAATTCCTAC 312
Qy      50 MevValAlaLeupheserAlaGlyLeuLeuLeuTyTyrAlaTyrLeuArgLysAsnAla 69
      313 GTGGTTCGGCTCTTAGTTCGATCCTTGGATATATTATGACCTTGAATCAACATG 372
Qy      70 TyrlleuIleValserIleLeuGlyPheGlyCysAlaIleGluLeuTyrIleSerLeu 89
      373 CCCCTACTCTATCAATTAATCTCCTTGGATGTTCTCATTTGAGACTATCTACGTTGGTTTC 432
Qy      90 PheLeuPheTyrAlaProArgLysSerLysIlePheThr---GlyTyrLeuMetLeuLeu 108
      433 TACTTTTCTTACGACCAAGAAAGGCGGCTCATCTATAAAAGCTCATATTATCA 492
Qy      109 GluLeuGlyAlaLeuGlyMetValMetProIleThrTyrLeuLeuAlaGlyLysSerHis 128
      493 GTGGTTCGTGGATTTGGTGCATATGCTGATCTGAATTTCTATTCAAAAGAGTGCTT 552
Qy      129 ArgValMetIleValGlyTyrPileCysAlaAlaIleAsnValAlaValPheAlaPro 148
      553 CGTGACAAATGTTGGATGGATTTGCTTATTTCTCTGTGTGTGATTTGTTGCTGCC 612
Qy      149 LeuSerIleLeuArgGlnValIleLeuThrLysSerValGluPheMetProPheThrLeu 168
      613 TTAGGCTTTGTGACACAGATCATCAAGCAGAGTGTGAAATCATGCACTTCTCTTA 672
Qy      169 SerLeuPheLeuThrLeuCysAlaThrMetTyrPhePheTyrGlyPhePheLysAsp 188
      673 TCAGTTTCTTCACATTATAGCGCTGTTATGCGTTCTTCTTATGCTCTTCTACTAAAGAT 732
Qy      189 PheTyrIleAlaPheProAsnIleLeuGlyPheLeuPheGlyIleValGluMetLeuLeu 208
      733 ATTAACATTTGCTGCCAAATGTTTGGGCTTCATCTTTGAAATATCTCCAAATTTATCTC 792
Qy      209 TyrPheAlaTyr---LysAspSerLysArgIleAspArgLysSerAspProValArg 227
      793 TATCATATATACACAAAGAAAGAGAGATCCATCTATAAGAACCAAACTTCCAGAGATA 852
Qy      228 GluAlaThrLysSerLysGluGlyValGluIleIleIleAsnIleGluAspAspAsnSer 247
      853 CAAGAAGCT-----GAAGTATTTGAAAGATGAGATGAAATGAAATGCA 894
Qy      248 AspAsnAlaLeuGlnSerMetGluLysAsp 257
      895 AATTAAGAAGCTTCCAGAACTCACCAAGAA 924
Db
RESULT 7
AKI03266 1562 bp mRNA linear PLN 24-JUL-2003
LOCUS DEFINITION
Oryza sativa (japonica cultivar-group) cDNA clone:J033124G19, full
insert sequence.
ACCESSION AKI03266
VERSION AKI03266.1 GI:32988475
KEYWORDS FLI_CDNA; CAP trapper.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
          Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
          Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
          Euphorbiales; Oryzaceae; Oryza.
REFERENCE
1 The Rice Full-length cDNA Consortium, National Institute of
  Agrobiological Sciences Rice Full-length cDNA Project Team;
  Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K.,
  Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I.,
  Kojima,K., Namiki,T., Ohneda,E., Yahagi,W., Yabagi,M., Suzuki,K., Li,C.,
  Ohtsuki,K., Shishiki,T., Foundation of Advancement of International
  Science Genome Sequencing & Analysis Group; Ootomo,Y., Murakami,K.,
  Iida,Y., Sugano,S., Fujimura,T., Suzuki,Y., Tsunoda,Y.,
  Narikawa,R., Kodama,T., Masuda,H., Kobayashi,M., Xie,Q., Lu,M.,
  Katsukawa,R., Sugiyama,A., Mizuno,K., Yokomizo,S., Niihara,J.,
  Ikeda,R., Ishibiki,J., Kawamata,M., Yoshimura,A., Miura,J.,
  Kusumegi,T., Oka,M., Ryu,R., Ueda,M., Matsubara,K., RIKEN;
  Kawai,J., Carninci,P., Adachi,J., Aizawa,K., Arikawa,T., Fukuda,S.,

```

Hara,A., Hashidume,W., Hayatsu,N., Imotani,K., Ishii,Y., Itoh,M.,  
 Kagawa,I., Kondo,S., Konno,H., Miyazaki,A., Otsu,N., Ota,Y.,  
 Saito,R., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T.,  
 Yoshino,M. and Hayashizaki,Y.  
 Collection, mapping, and annotation of over 28,000 cDNA clones from  
 japonica rice  
 Science 301 (5631), 376-379 (2003)

## JOURNAL MEDLINE PUBMED

## REFERENCE

## AUTHORS

TITLE  
 JOURNAL  
 COMMENT  
 URL: http://cdna01.dna.affrc.go.jp/cDNA/  
 NIAS Rice Full-length cDNA Project Team: Kikuchi,S., Satoh,K.,  
 Nagata,T., Kawagashira,N., Doi,K., Kishimoto,N., Yazaki,J.,  
 Ishikawa,M., Yamada,H., Ooka,H., Hotta,I., Kojima,K., Namiki,T.,  
 Ohneda,E., Yahagi,W., Suzuki,K., Li,C., Ohtsuki,K., Shishiki,T., and  
 Yamamoto,M.  
 FAIS Genome Sequencing & Analysis Group: Ootomo,Y., Iida,Y.,  
 Fujimura,T., Ikeda,R., Ishibiki,J., Kawamata,M., Kobayashi,M.,  
 Kodama,T., Kurosaki,T., Kusumegi,T., Lu,M., Masuda,H., Miura,J.,  
 Mizuno,K., Narikawa,R., Niihara,J., Oka,M., Ryu,R., Sugano,S.,  
 Sugiyama,A., Suzuki,Y., Tsunoda,Y., Ueda,M., Xie,Q., Yokomizo,S.,  
 Yoshimura,A., Matsubara,K. and Murakami,K.  
 Genome Exploration Research Group in Riken: Adachi,J., Aizawa,K.,  
 and Genome Science Laboratory in Riken: Adachi,J., Aizawa,K.,  
 Akimura,T., Arikawa,T., Carninci,P., Fukuda,S., Hangaki,T.,  
 Hara,A., Hashidume,W., Hayashida,K., Hayatsu,N., Hirooka,T.,  
 Hirakawa,T., Horii,F., Iida,Y., Imamura,K., Imotani,K., Ishii,Y.,  
 Itoh,M., Kagawa,I., Kanagawa,S., Katoh,H., Kawai,J.,  
 Kishikawa-Hirozane,T., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,  
 Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,  
 Nakamura,M., Nishii,K., Nomura,K., Numasaki,R., Ohno,M., Otsu,N.,  
 Ota,Y., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H.,  
 Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T.,  
 Sogabe,Y., Sato,K., Shibata,K., Shihagawa,A., Shiraki,T.,  
 Sogabe,Y., Tagami,M., Tagami-Takeda,Y., Tagawa,A., Takahashi,F.,  
 Takaku-Akashi,S., Tanaka,T., Tomaru,A., Toya,T., Tsunoda,Y.,  
 Ueda,M., Waki,K., Xie,Q., Yahagi,W., Yamada,H., Yamamoto,M.,  
 Yasunishi,A., Yazaki,J., Yokomizo,S. and  
 Yoshimura,A.  
 Direct Submission  
 Submitted (27-AUG-2002) Shoshi Kikuchi, National Institute of  
 Agrobiological Sciences, Department of Molecular Genetics, Head of  
 Laboratory of Gene Expression: 2-1-2 Kamondai, Tsukuba, Ibaraki  
 305-8602, Japan (E-mail: kikuchi@ias.affrc.go.jp,  
 Tel:81-29-838-7007, Fax:81-29-838-7007)  
 This clone is one of the 28K full-length cDNA clones from japonica  
 rice.

## FEATURES

## SOURCE

1. 1562  
 location/Qualifiers  
 /organism="Oryza sativa (japonica cultivar-group)"  
 /mol\_type="mRNA"  
 /cultivar="Nipponbare"  
 /db\_xref="taxon:39947"  
 /clone="J033124G19"

## ORIGIN

## Alignment Scores:

Pred. No.: 1,7e-50 Length: 1562  
 Score: 647.00 Matches: 116  
 Percent Similarity: 73.48% Conservative: 53  
 Best Local Similarity: 50.43% Mismatches: 55  
 Query Match: 47.82% Indels: 6  
 Gaps: 2

US-09-743-885a-1 (1-265) x AK103266 (1-1562)

QY 10 SerPheIlePheLeuGlyAsnIleValSerPheMetValPheLeuAlaProVal 29  
 Db 231 GCCTTACCTTGGACATCTTAAGTAATTAATCACTAATGATTCCTTGCCACATG 230  
 QY 30 ProThrPheTyrIleTyrIleValGlySerSerGluGlyTyrGlnAlaIleProTyr 49  
 Db 291 CCGACGCTTACCGCGGTACCGGAGAGAGATCGACGAGGGGTTCCAGTCAGCCCTAC 350  
 QY 50 MetValAlaLeuPheSerAlaGlyLeuLeuTyrTyrAlaTyrIleuArglyAsnAla 69  
 Db 351 GTGTGACGCTTACGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATG 410  
 QY 70 TyrLeuIleValSerIleLeuGlyPheGlyCysAlaIleGluLeuThrTyrIleSerLeu 89  
 Db 411 GACCTGCTGCTACACATCATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 470  
 QY 90 PheLeuPheTyrAlaProArgIleSerIlePheThrGlyTyrIleuMetLeu--Leu 108  
 Db 471 TACCTGCGCTACCGCCCAAGACGACGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 530  
 QY 109 GluLeuGlyAlaLeuGlyMetValMetProIleThrTyrIleuLeuAlaGluGlySerHis 128  
 Db 531 AACATGCGCTCTTGGCGCTGATGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 590  
 QY 129 ArgValMetIleValGlyTyrPileCysAlaIleLeuValAlaValPheAlaAlaPro 148  
 Db 591 CGGCTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 650  
 QY 149 LeuSerIleMetArgIleValIleLeuThrIleSerValGluPheMetProPheThrLeu 168  
 Db 651 TTGAGCATCATCAGCGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 710  
 QY 169 SerLeuPheLeuThrLeuCysAlaThrMetTrpPhePheTyrGlyPhePheValAsp 188  
 Db 711 TCCCTTCTTCTGCTGCTTACCGCGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 770  
 QY 189 PheTyrIleAlaPheProAsnIleLeuGlyPheLeuPheGlyIleValGlnMetLeuLeu 208  
 Db 771 GTGTTCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 830  
 QY 209 TyrPheValTyrIleAspSerIleValArgIleAspAluIleAspIleuSerAspProValArgIle 228  
 Db 831 TACATGGCGTACCGGACGAGAGCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 890  
 QY 229 AlaThrIleSerIleGluGlyValGlnIle 238  
 Db 891 GCC-----GGGCTGGAGATC 905

RESULT 8  
 AK101913  
 LOCUS  
 DEFINITION  
 Oryza sativa (japonica cultivar-group) cDNA clone:J033071H09, full  
 insert sequence.  
 AK101913  
 VERSION  
 AK101913.1 GI:32987122  
 KEYWORDS  
 FLI CDNA; CAP trapper.  
 SOURCE  
 Oryza sativa (japonica cultivar-group)  
 Oryza sativa (japonica cultivar-group)  
 Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 Ehrhartoideae; Oryzaceae; Oryza.  
 ORGANISM  
 Oryza sativa (japonica cultivar-group)  
 Oryza sativa (japonica cultivar-group)  
 Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 Ehrhartoideae; Oryzaceae; Oryza.  
 REFERENCE  
 1

## AUTHORS

The Rice Full-length cDNA Consortium, National Institute of  
 Agricultural Sciences Rice Full-length cDNA Project Team,  
 Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K.,  
 Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I.,  
 Kojima, K., Namiki, T., Ohneda, E., Yahagi, S., Suzuki, K., Li, C.,  
 Ohtsuki, K., Shishiki, T., Foundation of Advancement of International  
 Science Genome Sequencing & Analysis Group; Otsu, Y., Murakami, K.,  
 Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y.,  
 Kurosaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M.,  
 Naitaka, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Nishikura, J.,  
 Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, J.,  
 Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN,  
 Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S.,  
 Hara, A., Hoshidume, W., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M.,  
 Kagawa, I., Kondo, S., Kono, H., Miyazaki, A., Oosato, N., Ota, Y.,  
 Saito, R., Sasaki, D., Sato, K., Shibata, K., Shingawa, A., Shiraki, T.,  
 Yoshino, M., and Hayashizaki, Y.  
 Collection, mapping, and annotation of over 28,000 cDNA clones from  
 japonica rice  
 Science 301 (5631), 376-379 (2003)

JOURNAL  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORS

2 (bases 1 to 1494)  
 12865764  
 Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K.,  
 Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W.,  
 Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hiroake, T.,  
 Hori, F., Hotta, I., Iida, J., Iida, Y., Ikeda, R., Imamura, K.,  
 Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I.,  
 Kanagawa, S., Katoh, H., Kawagashira, N., Kawai, J., Kawamata, M.,  
 Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M.,  
 Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Kono, H., Kouda, M.,  
 Koya, S., Kurihara, C., Kurosaki, T., Kusumegi, T., Li, C., Lu, M.,  
 Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A.,  
 Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, M.,  
 Namiki, T., Naitaka, R., Nishikura, J., Nishikura, K., Nomura, K.,  
 Numasaki, R., Ohneda, E., Ohno, M., Ohtsuki, K., Oka, M., Ooka, H.,  
 Oosato, N., Ota, Y., Otsu, Y., Ryu, R., Saitoh, H., Sakai, C., Sakai, K.,  
 Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K.,  
 Shingawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S.,  
 Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y.,  
 Tagawa, A., Takehashi, F., Takaku-Akashita, S., Tanaka, T., Tomaru, A.,  
 Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, S.,  
 Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S., and  
 Yoshimura, A.  
 Direct Submission  
 Submitted (27-AUG-2002) Shoshi Kikuchi, National Institute of  
 Agricultural Sciences, Department of Molecular Genetics, Head of  
 Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki  
 305-8602, Japan (E-mail: kikuchi@nias.affrc.go.jp,  
 Tel: 81-29-838-7007, Fax: 81-29-838-7007)  
 This clone is one of the 28k full-length cDNA clones from japonica  
 rice.  
 URL: <http://cdna01.dna.affrc.go.jp/cdna/>  
 NIAS Rice Full-length cDNA Project Team: Kikuchi, S., Satoh, K.,  
 Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J.,  
 Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T.,  
 Ohneda, E., Yahagi, S., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T., and  
 Yamamoto, M.  
 FAIS Genome Sequencing & Analysis Group: Otsu, Y., Iida, Y.,  
 Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M.,  
 Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J.,  
 Mizuno, K., Naitaka, R., Nishikura, J., Oka, M., Ryu, R., Sugano, S.,  
 Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S.,  
 Yoshimura, A., Matsubara, K., and Murakami, K.  
 Genome Exploration Research Group in Riken Genomic Sciences Center  
 and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K.,  
 Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T.,  
 Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiroake, T.,  
 Hiroake, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y.,  
 Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J.,  
 Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Kono, H., Kouda, M.,  
 Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,  
 Nakamura, M., Naitaka, R., Nomura, K., Numasaki, R., Ohno, M., Oosato, N.,

## COMMENT

Ota, Y., Satoh, H., Sakai, K., Sakazume, N., Sano, H.,  
 Sasaki, D., Sato, K., Shibata, K., Shingagawa, A., Shireki, T.,  
 Sogabe, Y., Tagami, M., Tagami, T., Takada, Y., Tagawa, A., Takahashi, F.,  
 Takakura-Kahara, S., Tanaka, T., Tomaru, A., Toyota, T., Waki, K.,  
 Tsunashima, A., and Hayasizaki, Y.

FEATURES	Location/Qualifiers
source	1. .1494

## ORIGIN

### Alignment Scores:

Pred. No.:	1,8e-50	length:	149
Score:	64e.50	Matches:	119
Percent Similarity:	72.38	Conservative:	54
Best Local Similarity:	49.794	Mismatches:	55
Query Match:	47.784	Indels:	11
DB:	8	Gaps:	2

US-09-743-885A-1 (1-265) X AK101913 (1-1494)

Oy	10	SerpHellephgelyleuenglYsaniIeValsePmeMeValPheleuAlProVal	29
Db	260	GCCCTGCCCTTTGGTCTCTCTAGGCAACATCTCTTCATAGACTTACTGCCCCACTG	315
Oy	30	ProThrPheTYrLYsIleTYrLYsArgLYsSerSerGluGYrYrGlnAlaIleProTYr	49
Db	320	CCGACGTTCTACAGAGATCTTACAAGACAAGTGCAGAGGGGGTTCATCGTACCTTAC	375
Oy	50	MetValAlaIeuhSeerIaGlyLeuLeuLeuTYrTYrAlaTYrLeuArgLYsaAsnAla	69
Db	380	GTGGTGGCGCTGTTCAGCGCGATGCTGTGATCTACTACGCGCTCTCAAGTCGACAG	435
Oy	70	TYrIleuIleValSerIleasnGlyPheGlyCYsAlaIleGluLeuThrTYrIleSerIleu	89
Db	440	TGCTCTCTATACATCAACTCACTCGGTGGCTGGTGCATGAGACATCTACTATCGCGTC	495
Oy	90	PheLeuPheTYrAlaProArgLYsSerLYsIlePheThrGlyTYrPLeuMetLeuLeu--	106
Db	500	TACCTGCTTACGCCCCCAAGAGGCGAAGATGTTACCGCGCAACTCTCTCTCTGCTC	555
Oy	109	GluLeuGlyAlaIeuhGlyMetValmetProIleThrTYrLeuAlaGluGlySerHis	126
Db	560	AACGTGGCGCTCTTGGCCCTCACTCCCTCACCTCTCTCTCCGCGGAGACGC	615
Oy	129	ArgValMetIleValGlyTYrPLeCYsAlaAlaIleasnValAlaValPheAlaAlaPro	146
Db	620	CGATCGTGGTTCCTGGTGGGTCTGCTGCTCTCTCCGTACAGGCTTGTGTGCGCCCC	675
Oy	149	LeuSerIleMetArgGlnAlaIleLYsThrLYsSerValGluPheMetProPheThrIleu	166
Db	680	CTTAGCATCTACAGCGCTGTGTGGCCACCAAGAGGTGAGATTATGCGTCTTCGCTC	735
Oy	169	SerIeuhPheLeuThrIleuCYsAlaIleMetCTrPhePheTYrGlyPhePheLYsAsp	186
Db	740	TCTTCTCTCCCTACATCAAGCGCGTGGTGTCTCTACAGGCTCTCATCAAGAC	795
Oy	189	PheTYrIleAlaPheProAsnIleIeugLYPheLeuPheGlyIleValGluMetLeuLeu	206
Db	800	AAATATGTGCGCTTCTCCCAACGTGCGGGCTTCTCTTCGGGTATCTCAATATGGGGCTG	855
Oy	209	TYrPheValTYrLYsAspSerLYsArgIleAspArgGluLYsSerAspProValArgGln	226
Db	860	TACGCCATGTACAGAACTG-----ACGCCAAG	885
Oy	229	AlaThrLYsSerLYsGluGlyValGluIleIleIleasnIleGluAspAspAsnSer	247
Db	890	GCGGTGTGTACAGAGAGGTGCGAGCGGACGCGACCGAGGAGACAGACACTCC	946

AX506568	879 bp	DNA	linear	PAT 27-SEP-2002
LOCUS				
DEFINITION	Sequence	1263 from Patent WO0216655.		

**SOURCE** *Arabidopsis thaliana* (thale cress)

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
1  
Harper, J. F., Kreps, J., Wang, X., and Zhu, T.  
Stress-regulated genes of plants, transgenic plants containing  
same, and methods of use  
Patent: WO 021655-A 1263 28-FEB-2002;

## FEATURES

Source

FEATURES	Location/Qualifiers
source	1..879
	/organism="Arabidopsis thaliana"
	/mol_type="unassigned DNA"
	/db_xref="taxon:3702"

**ORIGIN**

### Alignment Scores:

Pred. No.:	1.3e-48	Length:	87
Score:	623.50	Matches:	11
Percent Similarity:	65.31%	Conservative:	58
Best Local Similarity:	43.91%	Mismatches:	63
Query Match:	46.08%	Indels:	31
DB:	6	Gaps:	4

US-09-743-885A-1 (1-265) X AX506568 (1-879)

Qy 1 LeuSerPheIlePheGlyLeuLeuGlyLysnIleValSerPheMetValPheLeuAlaPro 28  
Db 28 CTGCTTTATCTTGCGCATCTTAGCAACGGATATCTCTGTGATATCTTGCTTCA 87  
Qy 29 ValProThrPheTyrlsIleTyrlsAArgLysSerSerGluGlyTyrlGlnAlaIlePro 48  
Db 88 GTGCCAACTTTTATATGAATATATACAAAGAAATTCAGCGAAAGTTTCCAGTGCCTACG 147  
Qy 49 TyrMetValAlaLeuPheSerAlaGlyLeuLeuLeuTyrlTyrlAlaTyrlLeuAArgLysAsn 68  
Db 148 TACCAAGTGTCCCTATTAGCTGATGCAATGCTATAGGCTCTATACGATTGATTAAAGAAAC 207  
Qy 69 AlaTyrlLeuIleValSerIleAsnGlyPheGlyCysAlaIleGluLeuThrTyrlLeuSer 88  
Db 208 GCTTTTCTCTATTAACATCACTCACTCTTTGGCTGCGTCGGAGAGACTCTCTCATATGCC 267  
Qy 89 LeuPheLeuPheTyrlAlaProArgLysSerLysIlePheThrGlyTyrlPheMetLeu-- 107  
Db 268 ATGTTCTTGCTTACGCCACCGAGGAGAAAGATATGCGCTATGAAAGTTGTTTCATGCA 327  
Qy 108 LeuGluLeuGlyAlaLeuGlyLysMetValMetProIleThrTyrlLeuAlaGluGlySer 127  
Db 328 ATAAACGTGGCTTCTTCTCGTTGATCTTAATNGGTAACACATTTTCGGGTTAAACCTCT 387  
Qy 128 His--ArgValMetIleValGlyTyrlPheCysAlaIleAlaIleAsnValAlaValPheAla 146  
Db 388 CCCCTCAAGTCTGTGACTGCTGCGGATGTTGTGTGCATTCTGTGTTTCGTTTCGCT 447  
Qy 147 AlaProLeuSerIleMetArgGlnValIleLysThyrlLysSerValGluPheMetProPhe 166  
Db 448 GCCCTCTATATGATGCTGCTCGTGTGTAAAGCAAAAGATGGAGTGCATAGCCCTTC 507  
Qy 167 ThrLeuSerLeuPheLeuThrLeuGlyAlaIleMetTrpPhePheTyrlGlyPhePheLys 186  
Db 508 ACCCTTTCTTCTCTCATATAGCCCGGTATGCGTTCCGTATAGGTTTATTCCTC 567  
Qy 187 LysAspPheTyrlIleAlaPheProAsnIleLeuGlyPheLeuPheGlyIleValGlnMet 206

Db 568 AATGACATATGATGATGATTCACAAAGTGTGGATTCGTAAGGCGTGTCAATG 627

Qy 207 LeuLeuTyPheValTyrIleYsSerIleValArg----- 217

Db 628 GTTGTACTGTTGTTACGAACTCAATGAGAAACGAGAAATTAATTCGTGAGAA 667

Qy 218 -----11e 218

Db 688 CAACAATTAGAGTATGTCGTGATGAGTCCGTAGTGTGTCGGAAGTCACCCAGTT 747

Qy 219 AaPaPaGluIySerAaPaProValArgGluAlaThrIySerIleGluIyValGluIle 238

Db 748 GTGACCGAATCGGTGATCCACCTCTCTGAGCCGTT---CATGATGAGATGTGCAAA 804

Qy 239 llelleaenlleGluAaPaAenSerAaPaen 249

Db 805 GTTACTAAGTAGAGAGCCGTCATTTGAAAC 837

RESULT 10

AY113934 910 bp mRNA linear PLN 18-SEP-2002

LOCUS Arabidopsis thaliana putative senescence-associated protein SAG29

DEFINITION (At5g13170) mRNA, complete cds.

ACCESSION AY113934

VERSION AY113934.1 GI:21281009

KEYWORDS FLI CDNA.

SOURCE Arabidopsis thaliana (thale cress)

ORGANISM Arabidopsis thaliana

REFERENCE 1 (bases 1 to 910)

AUTHORS Yamada,K., Banb,J., Chan,M.M., Chang,C.H., Chang,E., Dale,J.M., Deng,J.M., Goldsmith,A.D., Lee,J.M., Onodera,C.S., Quach,H.L., Tang,C.C., Toriumi,M., Wu,H.C., Yamamura,Y., Yu,G., Bowser,L., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Lam,B., Lin,J., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shinn,P., Southwick,A., Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.

JOURNAL Arabidopsis Open Reading Frame (ORF) Clones

REFERENCE 2 (bases 1 to 910)

AUTHORS Yamada,K., Banb,J., Chan,M.M., Chang,C.H., Chang,E., Dale,J.M., Deng,J.M., Goldsmith,A.D., Lee,J.M., Onodera,C.S., Quach,H.L., Tang,C.C., Toriumi,M., Wu,H.C., Yamamura,Y., Yu,G., Bowser,L., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Lam,B., Lin,J., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shinn,P., Southwick,A., Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.

JOURNAL Direct Submission

COMMENT Submitted (16-MAY-2002) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA

COMMENT The RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAPL cDNAs (RAPL cDNA: 'RIKEN Arabidopsis Full-length cDNA'): Seki,M., Narusaka,M., Ishida,J., Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J., Hayashizaki,Y. and Shinozaki,K.

COMMENT The Salk, Stanford, PGEC (SSP) Consortium members constructed and sequenced the PUNI (ORF) clones using the RAPL cDNAs: Yamada,K., Banb,J., Chan,M.M., Chang,C.H., Chang,E., Dale,J.M., Deng,J.M., Goldsmith,A.D., Lee,J.M., Onodera,C.S., Quach,H.L., Tang,C.C., Toriumi,M., Wu,H.C., Yamamura,Y., Yu,G., Bowser,L., Cheuk,R., Jones,T., Karlin-Neumann,G., Kim,C., Lam,B., Lin,J., Miranda,M., Nguyen,M., Palm,C.J., Shinn,P., Southwick,A., Davis,R.W., Ecker,J.R. and Theologis,A.

COMMENT Yamada,K. (SSP/PGEC) and Seki,M. (RIKEN GSC) contributed equally to this work. Shinozaki,K. (RIKEN GSC) and Theologis,A. (SSP/PGEC) contributed equally to this work as PIs.

Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank.

FEATURES

source

1..910

/organism="Arabidopsis thaliana"

/mol\_type="mRNA"

/db\_xref="taxon:3702"

/chromosome="5"

/clone="U10278"

/ecotype="Columbia"

/note="This clone is in PUNI 51."

1..910

/gene="At5g13170"

1..879

/gene="At5g13170"

/note="contains prokaryotic membrane lipoprotein lipid attachment site A447-57"

/codon\_start=1

/evidence=experimental

/product="putative senescence-associated protein SAG29"

/protein\_id="AA044982.1"

/db\_xref="GI:21281010"

/translation="MGVMTNHPFLAFIIGILNVISFLVFLAPVPTFYIYRKSTES FQSLPYQVSLFSCMLLYVALIKDAFLITINSFCCVETIYIAFPYATREKIS AKMLFIANNVAFSLIATHTHFVKTPLPQVGLMI CVAISVSPADLMTVARIK TKSVEVPTLSPFLTISAVMPFVGLFNDICAIENYVPLGLQWLVLYVENS NKREKINSFQOLKSIYVMSPLGVSEHVPVTESVDPISLVHHDLSKVTKEEPS IENGKYEATRETEY"

880..910

/gene="At5g13170"

ORIGIN

3' UTR

Alignment Scores:

Pred. No.: 1,35e-48 Length: 910

Score: 623.50 Matches: 119

Percent Similarity: 65.31% Conservative: 58

Best Local Similarity: 43.91% Mismatches: 63

Query Match: 46.08% Indels: 31

DB: 8 Gaps: 4

US-09-743-885A-1 (1-265) x AY113934 (1-910)

Qy 9 LeuSerPheIlePheGlyLeuLeuGlyAanIleValSerPheMetValPheLeuAlaPro 28

Db 28 CTCGCTTTATCTTCGGCATCTTAGAAGCGATATCTCTGTGATTCCTCGCTCCA 87

Qy 29 ValProThyPheTyrIleTyrIleTyrIleTyrIleTyrIleTyrIleTyrIleTyrIle 48

Db 88 GTCCCAACTTTTATAGATATATACAGAGAAATCGACGGAAGTTCCAGTCGTACCG 147

Qy 49 TyrMetValAlaLeuPheSerAlaGlyLeuLeuTyrTyrIleTyrIleTyrIleTyrIle 68

Db 148 TACCAAGTGCCTATTATGCTGCATGCTATGCTTACTACCATTTGATTAAGAAAGAC 207

Qy 69 AlaTyrLeuIleValSerIleLeuGlyPheGlyCysAlaIleGluLeuThrTyrIleSer 88

Db 208 GCTTTCTCTTAATTAACATCAACTCTTGGCTGCTGTGAAGACTCTCTACATAGCC 267

Qy 89 LeuPheLeuPheTyrIleAlaProArgIySerIyIlePheThrGlyTyrPheMetLeu--- 107

Db 268 ATGTTCTTCGCTTAACCAACAGGAGAAAGATATCGGCTATGAAGTTGTCATAGCA 327

Qy 108 LeuGluLeuGlyAlaLeuGlyMetValIlePheThrTyrIleLeuAlaGluIySer 127

Db 328 ATGAAGTGGCTTCTTCGTTGATTCATAGGTACACATTTCGTGTTAAACTCCT 367

Qy 128 His---ArgValMetIleValGlyTyrIleCysAlaAlaIleLeuValAlaValPheAla 146

Db 388 CCCCTCAAGTCTCTGATCTGCGCTGATTTGCTTCCATTTCTGTTTCGTTTCGCT 447

Qy 147 AlaProLeuSerIleMetArgGluValIleLeuThrIySerValGluPheMetProPhe 166

Db 448 GCCCTCTAATGATCGTGCTCGTGTGATTAAGACAAAGTGTGAGTGCATGACCTTC 507

QY 167 ThrLeuSerLeuPheLeuThrLeuCysAlaThrMetTrpPhePheTygIlyPhePheLys 186  
Db 508 ACGCTTCTTCTTCTCTCCTACATATAGCGCCGCTTATGATGCTTATGCTTATGCTTCTC 567  
QY 187 LysAspPheTyrlleAlaPheProAnlleuGllyPheLeuPheGlylleValGlnMet 206  
Db 568 AATGACATATGACATAGCGATTCCAAACGTGTGGATTCGTATAGGCGCTGTTCAAATG 627  
QY 207 LeuLeuTyrlPheValTyrlLysAspSerLysArg----- 217  
Db 628 GTTTTGTACTGCTGCTTACAGAACTCAATAGAAACCAAGAAATTAATTCGTCAAA 687  
QY 218 -----Ile 218  
Db 688 CAACACTTAAGATATGTCGTATGATGATCGCTTAAAGTGTGCGAAGTGCACCCAGTT 747  
QY 219 AspAspGluLysSerAspProValArgGluAlaThrLysSerLysGluGlyValGlnIle 238  
Db 748 GTGACGGAATCGTGTGACCCACTCTGTGAACCGCTT--CATCATGAGGATCTGCCAA 804  
QY 239 llelleAenlleGluAspAspAspSerAspAsn 249  
Db 805 GTTACTAAAGTGTGAGGAGCCCTCAATTGAAC 837  
RESULT 11  
AY087836 1269 bp mRNA linear PLN 14-APR-2003  
LOCUS Arabidopsis thaliana clone 38843 mRNA, complete sequence.  
DEFINITION AY087836  
ACCESSION AY087836  
VERSION AY087836.1 GI:21406587  
KEYWORDS Full cDNA.  
SOURCE Arabidopsis thaliana (thale cress)  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
REFERENCE 1 (bases 1 to 1269)  
Haas,B.J., Volkovsky,N., Town,C.D., Troukhan,M., Alexandrov,N.,  
Feldmann,K.A., Flavell,R.B., White,O. and Salzberg,S.L.  
Full-length messenger RNA sequences greatly improve genome  
annotation  
Genome Biol. 3 (6), RESEARCH0029 (2002)  
JOURNAL MEDLINE 22088475  
PUBMED 12093376  
REFERENCE 2 (bases 1 to 1269)  
Brover,V., Troukhan,M., Alexandrov,N., Lu,Y.-P., Flavell,R. and  
Feldmann,K.  
Full-length cDNA from Arabidopsis thaliana  
Unpublished  
3 (bases 1 to 1269)  
Brover,V., Troukhan,M., Alexandrov,N., Lu,Y.-P., Flavell,R. and  
Feldmann,K.  
Direct Submission  
Submitted (11-MAR-2002) Ceres, Inc, 3007 Malibu Canyon Road,  
Malibu, CA 90265, USA  
COMMENT  
TITLE This clone sequence is one of 5,000 Ceres full-length cDNAs made  
JOURNAL available to TIGR and Genbank. The following quality assessment of  
REFERENCE this set was done by comparison with known proteins: two percent  
AUTHORS of the clones are estimated to be 5'-truncated; less than one percent  
are 3'-truncated; approximately two percent represent alternative  
splice variants, including unspliced introns and spliced exons; one  
percent may contain premature stop codons; five percent may have  
frame shifts in a coding region. A sequence is considered to be  
5'-truncated if it lacks the translation initiation start (ATG). A  
sequence is considered to be 3'-truncated if it lacks the  
C-terminal end of the encoded protein. Please note that these cDNA  
sequences are derived from the Ws or later ecotypes and therefore  
may contain polymorphisms when compared to sequences from Col-0.  
Genet carried out the library production and sequencing of the  
full-length clones. Ceres, Inc. carried out the clustering of the  
5' sequences, selection of clones, and sequence assembly.  
FEATURES  
Location/Qualifiers

source  
1. 1269  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/db\_xref="taxon:3702"  
/clone="38843"  
CDS  
136..1014  
/codon\_start=1  
/product="senescence-associated protein (SAG29)"  
/protein\_id="AA065389.1"  
/db\_xref="GI:21593422"  
/translation="MGVILNHPLAFPGILGVISFLVFLAVPVPFVRYKRKSTES  
FQSLPYQVSLFSCMLLYALIKKDAFLITINSRCVVEVTLIAMPFAATREKIS  
AKMLFTAMNVAFPFLIMVTHFVYKTPPLQVSLGVIQAISISVPAAPLMTYRITK  
TSKVEYMPFLSEFLTISAVMFPAYGLFINDICAIPLNVGFLGLQWLVLYRNS  
NEKPEKINSSEEQIKSIVVMSPLGVSEVHPVTVESVDPVLSAVHBDLSKVTVEBPS  
IENKGVCEVATREPETV"  
ORIGIN  
Alignment Scores:  
Pred. No.: 2e-48 Length: 1269  
Score: 623.50 Matches: 119  
Percent Similarity: 65.31% Conservative: 58  
Best Local Similarity: 43.91% Mismatches: 63  
Query Match: 46.08% Indels: 31  
DB: Gaps: 4  
US-09-743-885A-1 (1-265) x AY087836 (1-1269)  
QY 9 LeuSerPheLlePheGlyLeuLysAlaThrMetValPheLeuAlaPro 28  
Db 163 CTGCTTTATCTTGGCATCTTAGGAAACGTGATATCTTCTTGTATCTTGGCTTCCA 222  
QY 29 ValProThrPheTyrlLysSerLysArgLysSerSerGluGlyArgAlaIlePro 48  
Db 223 GTGCGAATCTTTATAGAAATATCAAGAGAAATTCAGCGAAAGTTCCAGTGCCTACCG 282  
QY 49 TyrlMetValAlaLeuPheSerAlaGlyLeuLeuLeuTyrlAlaTyrlLeuAlaGlyLysAsn 68  
Db 283 TACCAAGTGTGCGTATTTAGCTCATGCTCATGCTCTTACTACCATGATGAAGAAAGAC 342  
QY 69 AlaTyrlLeuLysAlaSerlleAenlleAenGlyPheGlyCysAlaIleGluLeuThrTyrlLysSer 88  
Db 343 GCTTTTCTCTTAATATACATCAACTCTTGTGGCTGCGTGGAGACTCTTACATAGCC 402  
QY 89 LeuPheLeuPheTyrlAlaProArgLysSerLysAlaPheThrGlyTrpLeuMetLeu--- 107  
Db 403 ATGTTCTTGCCTTACCGCACAGGAGAAAGATATCGGCTATGAACTTGTTCATAGCA 462  
QY 108 LeuGluLeuGlyAlaLeuGlyMetValMetProIleThrTyrlLeuLeuAlaGluGlySer 127  
Db 463 ATGACCGTTCCTCTTCTGCTGATTCATATGATACATTCGCGTTAAACCTCCT 522  
QY 128 His---ArgValMetlleValGlyTrpIleCysAlaIleAlaIleAsnValAlaIlePheAla 146  
Db 523 CCCCTCCAAAGTCTCTGACTCGGTGATTCGATTCGCTTCTGTTCTGTTTTCGCT 582  
QY 147 AlaProLeuSerlleMetArgGlnValIleLysThrLysSerValGluPheMetProPhe 166  
Db 583 GCCCTTATATGATCGCGCTCGATGATTAAGACAAAGGTGAGGATCAAGCCCTTC 642  
QY 167 ThrLeuSerLeuPheLeuThrLeuCysAlaThrMetTrpPhePheTygIlyPhePheLys 186  
Db 643 ACGCTTCTTCTTCTCCTACATATAGCGCCGCTTATGCTTATGCTTATGCTTCTC 702  
QY 187 LysAspPheTyrlleAlaPheProAnlleuGllyPheLeuPheGlylleValGlnMet 206  
Db 703 AATGACATATGACATAGCGATTCCAAACGTGTGGATTCGTACTAGGCGCTTGCAATG 762  
QY 207 LeuLeuTyrlPheValTyrlLysAspSerLysArg----- 217  
Db 763 GTTTTGTACTGCTGCTTACAGAACTCAATAGAAACCAAGAAATTAATTCGTCAAA 822  
QY 218 -----Ile 218

Db 823 GAACAACTTAAGATATTGTCGTATGATGATCCGTAGTGTGTGCGAAGTGCACCCAGTT 882  
Qy 219 AAGPAGGULysSerAPProValArgLValThrLysSerLysGluGlyValGluIle 238  
Db 883 GTGACCGAATCGGTGATCCACCTCTCTGTAAGCCGTT---CATCATGAGATCTGTCCAA 939  
Qy 239 IleIleAlaIleGluAlaPheAAsSerAAsPAsn 249  
Db 940 GTTACTTAAGATGAGAGAGCCGTCAATTGAAAC 972  
RESULT 12  
LOCUS AY045949  
DEFINITION Arabidopsis thaliana putative senescence-associated protein SAG29  
ACCESSION AY045949  
VERSION AY045949.1 GI:15028292  
KEYWORDS FLI CDNA.  
SOURCE Arabidopsis thaliana (thale cress)  
ORGANISM Arabidopsis thaliana  
REFERENCE 1 (bases 1 to 1291)  
AUTHORS Yamada,K., Liu,S.X., Sakano,H., Pham,P.K., Banh,J., Dale,J.M., Goldsmith,A.D., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Lam,B., Lin,J., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T., Sato,M., Seki,M., Shim,P., Southwick,A., Shinzaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.  
TITLE Arabidopsis Full Length cDNA Clones  
JOURNAL Unpublished  
AUTHORS 2 (bases 1 to 1291)  
Yamada,K., Liu,S.X., Pham,P.K., Banh,J., Dale,J.M., Goldsmith,A.D., Jiang,P.X., Lee,J.M., Onodera,C.S., Quach,H.L., Tang,C.C., Toriumi,M., Yamamura,Y., Yu,G., Yu,S., Bowser,L., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Koeseema,E., Lam,B., Lin,J., Meyers,M.C., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T., Sato,M., Seki,M., Shim,P., Southwick,A., Shinzaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.  
COMMENT Submitted (10-JUL-2001) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA  
RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA: 'RIKEN Arabidopsis Full-length cDNA'): Seki,M., Narusaka,M., Ishida,J., Sato,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J., Hayashizaki,Y. and Shinzaki,K.  
The Salk, Stanford, PGSC (SSP) Consortium members carried out the sequencing and annotation of the RAFL cDNAs: Yamada,K., Liu,S.X., Pham,P.K., Banh,J., Dale,J.M., Goldsmith,A.D., Jiang,P.X., Lee,J.M., Onodera,C.S., Quach,H.L., Tang,C.C., Toriumi,M., Yamamura,Y., Yu,G., Yu,S., Bowser,L., Chen,H., Cheuk,R., Jones,T., Karlin-Neumann,G., Kim,C., Koeseema,E., Lam,B., Lin,J., Meyers,M.C., Miranda,M., Nguyen,M., Palm,C.J., Shim,P., Southwick,A., Tracy,S.E., Davis,R.W., Ecker,J.R. and Theologis,A.  
Yamada,K. (SSP/PGSC) and Seki,M. (RIKEN GSC) contributed equally to this work. Shinzaki,K. (RIKEN GSC) and Theologis,A. (SSP/PGSC) contributed equally to this work as PIs.  
FEATURES  
source  
Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank.  
location/Qualifiers  
1..1291  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/db\_xref="taxon:3702"  
/chromosome="5"

ORIGIN  
Alignment Scores:  
Pred. No.: 2,04e-48 Length: 1291  
Score: 623.50 Matches: 119  
Percent Similarity: 65.31% Conservative: 58  
Best Local Similarity: 43.91% Mismatches: 63  
Query Match: 46.08% Indels: 31  
Gaps: 4  
US-09-743-885a-1 (1-265) x AY045949 (1-1291)  
Qy 9 LeuSerPheIlePheGlyLeuGlyAAsIleValSerPheMetValPheLeuAlaPro 28  
Db 164 CTGGCTTTATCTTCGGCATCTTAGAAGCGGATATCTCTCTTATTCCTCGCTCA 223  
Qy 29 ValProThrPheTyrLysIleTyrLysArgLysSerSerGluGlyTyrGlnAlaIlePro 48  
Db 224 GTGCCAAGCTTTTATGATATATACAGAAATGACGGAAGTTCCAGTGCCTACCG 283  
Qy 49 TyrMetValAlaLeuPheSerAlaGlyLeuLeuTyrTyrAlaTyrLeuArgLysAsn 68  
Db 284 TACCAAGTCTCGCTATTATGCTGCATGCTATGCTCTACTACGCAATGATTAAGAAAGC 343  
Qy 69 AlaTyrLeuIleValSerIleAsnGlyPheGlyCysAlaIleGluLeuThrTyrIleSer 88  
Db 344 GCTTCTCTATTAACATCACTCTTGGCTGGTGTGGAGACTCTTCAATACATAGCC 403  
Qy 89 LeuPheLeuPheTyrAlaProArgLysSerLysIlePheThrGlyTyrPheMetLeu--- 107  
Db 404 ATGTTCTTGGCTTAGCCACCAAGGAGAAAGATATCGGCTATGAAAGTTTCATAGCA 463  
Qy 108 LeuGluLeuGlyAlaLeuGlyMetValMetProIleThrTyrLeuLeuAlaGluGlySer 127  
Db 464 ATGAAAGTCCCTTCTTCGCTGATTAAGCTTAACCAATTCCTGCTTAAGTCCCT 523  
Qy 128 His---ArgValMetIleValGlyTyrPheCysAlaAlaIleAsnValAlaValPheAla 146  
Db 524 CCCCTCAAGTCTCTACTCGGCTGGAATTTGTGTCATTTCTCTTCTGTTTCGCT 583  
Qy 147 AlaProLeuSerIleMetArgGlnValIleLysThrLysSerValGluPheMetProPhe 166  
/clone="RAFLO5-19-F21 (R10278)"  
/ecotype="Columbia"  
/note="This clone is in a modified pBluescript vector (lambda ZAP) as a XhoI/SacI insert."  
1..1291  
/gene="At5g13170"  
1..136  
/gene="At5g13170"  
84  
/gene="At5g13170"  
/note="not present in genomic sequence"  
137..1015  
/gene="At5g13170"  
/note="Contains Prokaryotic membrane lipid attachment site AA47-57"  
/codon\_start=1  
/evidence="experimental"  
/product="putative senescence-associated protein SAG29"  
/protein\_id="AAK76623.1"  
/db\_xref="GI:15028293"  
/translation="MGVMNHFLAFIPGILGNVISFLVFLAPVPTFRIYRKSTES  
FQSLPQVSLRESCMLMYALIKKAPFLITNSRGCVETIYIMFPYARREKIS  
AKKLFTAMNVAFFSLIMVTHFVATPPLQVSLGVCAISVFAALPLTVARVYK  
TQSVEMPTLSFPLTISAVMFAVGLFLNDICIAIPNVGFLGLQVLYLVKNS  
NEKPEKINSBQOLKSIIVMSPLGVSEVHPVTESVDPLSEAVHEDLSKTKVEPS  
IENGCVVATPEPTV"  
1016..1291  
/gene="At5g13170"  
1276  
/gene="At5g13170"  
/note="artifact within poly A tail"

Db	584	GGCCCTCAATCATGCTGGCTCGTGTGATTAAGACAAAGGTGAGATCATGCGCTTC	643
Qy	167	ThrluSeSLeuPheLeuThrluCySaIarhMeTrlPhePheTySlYpHePheLys	186
Db	644	ACGCTTCTCTTCTCCTCACTATMAAGCGCGTTATGTGGTTCGCTATGTTATTCCTC	703
Qy	187	LysAaPheTyTlLeAlaPheProkaniLeLeuGlyPheLeuPheGlyLlLeValGlnMet	206
Db	704	AATGACATPATGATGATGCGATTCCAAACGGGGTGGAGATTGTCATGAGGCGCTTGACAAATG	763
Qy	207	LeuLeuTyPheValTyLysAspSerLysA-rg-	217
Db	764	GTTTGTACTGCTTTACAGAACTCAATATAGAAAACAGAAAGATTATTCGTACAGA	823
Qy	218	-----	1le 218
Db	824	CAACAACTTAAGATATTGTCGTGATGATCGCTAGTGTGTCGGAGATGCACCAAGTT	883
Qy	219	AspaAPGluLysSerAspProValargLualatrlYsSerLysgluLylValGluIle	238
Db	884	GTGACGGATCGGTGACCCACTCTCTGAAGCGTT--CATCATGAGATCTGTCCAA	940
Qy	239	IleIleAsnIleGluAspaSpsanSerAspAn	249
Db	941	GTTACTAAAGTGGAGGAGCGCCATCATTTGAAAC	973
RESULT 13			
CQ805384		870 bp	DNA
LOCUS			linear
DEFINITION	Sequence 1795 from Patent WO2004035798.		PAT 10-MAY-200
ACCESSION	CQ805384		
VERSION	CQ805384.1	GI:47111326	
KEYWORDS			
SOURCE			
ORGANISM	Arabidopsis thaliana (thale cress)		
	Arabidopsis thaliana		
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;		
	rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.		
REFERENCE			
AUTHORS	Inze,D., de Veylder,L. and Vliethe,K.		
TITLE	Identification of novel e2f target genes and use thereof		
JOURNAL	Patent: WO 2004035798-A 1795 25-APR-2004;		
	CropDesign N.V. (BE)		
FEATURES			
source	Location/Qualifiers		
	1..870		
	/organism="Arabidopsis thaliana"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:3702"		
ORIGIN			
Alignment Scores:			
Pred. No.:	1..77e-48	Length:	870
Score:	622.00	Matches:	126
Percent Similarity:	65.40%	Conservative:	46
Best Local Similarity:	47.91%	Mismatches:	63
Query Match:	45.97%	Indels:	28
DB:	6	Gaps:	4
US-09-743-885A-1 (1-265) x CQ805384 (1-870)			
Qy	10	SerPheIlePheGlyLeuLeuGlyAsnIleValSerPheMetValPheLeuAlaProVal	29
Db	31	GCCTTGTCTTGGCTGCTCGGCAACCTTATCTCTTGGCCGTGTTCTATCTCTG	90
Qy	30	ProThrPheTyLysIleTyLysArgLysSerSerGluGlyTyrsGlnAlaIleProTy	49
Db	91	CCAAAGCTTCTAATAGGATTTGGAAAGAAACAACAGAAAGGTTTCAAGTATCTCTAT	150
Qy	50	MetValAlaLeuPheSerAlaGlyLeuLeuLeuTyTyTyAlaTyLysArgLysAsnAla	69
Db	151	GTTGTGGGCGCTTTCAGTGCAGCGCTTGGCTTTACTATAGCCACACAGAAAGAAATGTC	210
Qy	70	TyLysLeuIleValSerIleAsnGlyPheGlyCySaIaIleGluLeuThrTyLleSerLeu	89

Db		211	TTCTCCCTCGTAACCATTAAGCGCCTTGGTGGTTCATCAAGAACCATTCAATCTATG	270
OY		90	PheLeuPheTyAlaProArgLysSerIleValIlePheThr--GIyTrpLeuMetLeuLeu	108
Db		271	TTCTTGCCCTACGCCTCCCAAGCCAGCGCTCGGATGTGTACAAGAAAGTACTACTTATG	330
OY		109	GlUeUGlYalaleUGlyMetValMetProIIeThrTyLeuEuaIagUGlySerHis	128
Db		331	AAC TT TG AG AG ATT CT GT CG AT T CT C TC TC TT G CC AA TT CT G T T A A A A G G C C A C A	390
OY		129	ArgValMetIleValIGLyTrpIleCysAlaIalaIleasnValaIaValIPheaAlaPro	148
Db		391	CGTCTTAGAATTATCGAGAGGAATGTGTGTGGATTCCTGTGTGTGTGTTCGCTGCTCT	450
OY		149	LeuSerIleMetArgInValIIelValThrLysSerValGIuPheMetProPheThrLeu	168
Db		451	CTAAGCAATAATCACAGACGATATAAACAAGAAGTGTGGATGCATAGCCCTTAGCTTA	510
OY		169	SerLeuPheLeuThrLeuCysAlaIleMetTrpPhePheTyrgIYPhePheLYsAsp	188
Db		511	TCCTTAACCCCTTACCATCAGTGCATGCTCATATGGCTCCTTATGTGCTTGCTTCACAGAC	570
OY		189	PheTyrlleAlaPheProanilleUGlyPheLeuHegIyleValIGImetLeuLeu	208
Db		571	ATCTATGTGTGTTTCCGAATGTGCTGTGTGCTCTCGTGCACTCCAAAATGATATC	630
OY		209	TyrPheValTYrLys-----AspSerLYsArgLIeAspasp	220
Db		631	TACCTGTGCTACAAATACTGTAAAAAGCTGCCCGCATCTAGAGAGAAAGAAGTGAAGCT	690
OY		221	GIuLYs-----SerAspPro---SerpPro---	225
Db		691	GCTAGAGTTACCGAGGTGAGCCTCGATATGTTGAAGCTAGGACACAGTTTATCCCTGAG	750
OY		226	-----ValARGGIualatThrLYsSerLYsGIuGIyValGIuIlelleIleAsn	241
Db		751	CCAATCTCAGTGTGTCGACCAACAGATGTACTCGGAAATGATCGAAGGCTGAG	810
OY		242	IleGIuAsp	244
Db		811	ATTGAAGAT	819
RESULT 14				
AY078041	LOCUS	870 bp	mRNA	linear
AY078041	DEFINITION	Arabidopsis thaliana AT3g48740/18P19_250 mRNA, complete cds.		
AY078041	VERSION	AY078041.1 GI:18700263		
KEYWORDS	Fli CDNA.			
SOURCE	Arabidopsis thaliana (thale cress)			
ORGANISM	Arabidopsis thaliana			
REFERENCE	AUTHORS			
1	(bases 1 to 870)			
Cheuk,R., Chen,H., Kim,C.-J., Meyers,M.C., Banh,J., Bowser,L.,				
Carninci,P., Chang,E., Dale,J.M., Goldsmith,A.D., Hayaishiizaki,Y.,				
Ishida,T., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J.,				
Lam,B., Lee,J.M., Lin,J., Miranda,M., Narusaka,M., Nguyen,M.,				
Onodera,C.S., Palm,C.U., Quach,H.L., Sakurai,T., Satou,M., Seki,M.,				
Southwick,A., Tang,C.C., Toriumi,M., Wu,H.C., Yamada,K.,				
Yamanura,Y., Yu,G., Yu,S., Shinozaki,K., Davis,R.W., Theologis,A.				
and Ecker,J.R.				
Arabidopsis CDNA clones				
Unpublished				
2	(bases 1 to 870)			
Cheuk,R., Chen,H., Kim,C.-J., Meyers,M.C., Banh,J., Bowser,L.,				
Carninci,P., Chang,E., Dale,J.M., Goldsmith,A.D., Hayaishiizaki,Y.,				
Ishida,T., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J.,				
Lam,B., Lee,J.M., Lin,J., Miranda,M., Narusaka,M., Nguyen,M.,				
Onodera,C.S., Palm,C.U., Quach,H.L., Sakurai,T., Satou,M., Seki,M.,				
Southwick,A., Tang,C.C., Toriumi,M., Wu,H.C., Yamada,K.,				
Yamanura,Y., Yu,G., Yu,S., Shinozaki,K., Davis,R.W., Theologis,A.				
and Ecker,J.R.				
Arabidopsis CDNA clones				
Unpublished				





TITLE  
JOURNAL  
COMMENT

Shinozaki, K., Davis, R.W., Becker, J.R. and Theologis, A.  
Direct Submission  
Submitted (16-APR-2002) Plant Gene Expression Center, 800 Buchanan  
Street, Albany, CA 94710, USA  
The RIKEN Genomic Sciences Center (GSC) members carried out the  
collection and clustering of RAPL cDNAs (RAPL cDNA: 'RIKEN  
Arabidopsis Full-Length cDNA'): Seki, M., Narusaka, M., Ishida, J.,  
Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J.,  
Hayashizaki, Y. and Shinozaki, K.

The Salk, Stanford, PGEC (SSP) Consortium members constructed and  
sequenced the pUNI (ORF) clones using the RAPL cDNAs: Yamada, K.,  
Bain, J., Chan, M.M., Chang, C.H., Chang, E., Dale, J.M., Deng, J.M.,  
Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C.,  
Toriumi, M., Wu, H.C., Yamamura, Y., Yu, G., Bowser, L., Chen, H.,  
Cheuk, R., Jones, T., Karlin-Neumann, G., Kim, C., Lam, B., Lin, J.,  
Meyers, M.C., Miranda, M., Nguyen, M., Palm, C.J., Shinn, P.,  
Southwick, A., Davis, R.W., Becker, J.R. and Theologis, A.  
Yamada, K. (SSP/PGEC) and Seki, M. (RIKEN GSC) contributed equally  
to this work. Shinozaki, K. (RIKEN GSC) and Theologis, A. (SSP  
/PGEC) contributed equally to this work as PIs.

Annotation is based on the January 2002 version of the Arabidopsis  
genome submitted to Genbank.

FEATURES  
source

Location/Qualifiers

1..901  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/db\_xref="taxon:3702"  
/chromosome="3"  
/clone="U19389"  
/ecotype="Columbia"  
/note="This clone is in pUNI 51."

gene  
CDS  
1..870  
/gene="At3g48740"  
/codon\_start=1  
/evidence="experimental"  
/product="putative MTN3 protein"  
/protein\_id="AA02044.1"  
/db\_xref="GI:20465523"  
/translation="MSLFNTFWAFVFGGLGNLISPAVLSPPVTFYRIWKKTTEG  
FOSIPYVALFSATLMLYVAIOKDVPLVTINAFGFIETIYSMLAVAPKARML  
TVKLLMNFSGFCAILLCOPLVKGATRAKILIGICFVSVCVPAFLAFLIRIVIT  
RSVEMPEFSILITISAVIWLIGALKIIVYAPNVILGALGALMLIIVYVYKCK  
TSRHLGEVBAKLPVSLDMLKLTGVSSPEPISVVRANKTCGADRAREIDGQT  
PKHGKQSSAAAT"  
871..901  
/gene="At3g48740"

ORIGIN  
3' UTR  
/gene="At3g48740"

Alignment Scores:

Pred. No.: 1.84e-48 Length: 901  
Score: 622.00 Matches: 126  
Percent Similarity: 65.40% Conservative: 46  
Best Local Similarity: 47.91% Mismatches: 63  
Query Match: 45.97% Indels: 28  
DB: 8 Gaps: 4

US-09-743-885a-1 (1-265) x AY096594 (1-901)

Qy 10 SerPheHlePheGlyLeuLeuGlyAsnIleValSerPheMetValPheLeuAlaProVal 29  
Db 31 GCCTTGTGCTTGGCTTGGCTGCAACTTATCTCTTGGCTGCTTCTATCTCTG 90  
Qy 30 ProThrPheTyrIleTyrIleTyrIleTyrIleTyrIleTyrIleTyrIleTyrIleTyr 49  
Db 91 CCAAGCTTCTTAAAGATTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 150  
Qy 50 MetValAlaLeuPheSerAlaGlyLeuLeuLeuTyrTyrAlaTyrLeuArgIleAsnAla 69

Db 151 GTTGTGGCGCTTTCAGTGGAGCGCTTGGCTTACTATGCGACAGAGAGAGATGTC 210  
Qy 70 TyrIleuIleValSerIleAsnGlyPheGlyCysAlaIleGluLeuThrTyrIleSerLeu 89  
Db 211 TTCTCTCTGCTAACCATTAACGCTTGGTGTCTCATGCAACCATTCATCTCATG 270  
Qy 90 PheLeuPheTyrAlaProAlaGlySerIlePheThr---GlyTyrLeuMetLeuLeu 108  
Db 271 TTCTTGGCTTACCGCTCCAGCGACGCTCGATGTGACAGTGAAGATGCTTCTTANG 330  
Qy 109 GluLeuGlyAlaLeuGlyMetValMetProIleThrTyrLeuLeuAlaGluGlySerHis 128  
Db 331 AACTTGGAGAGATTCGTGTGATCTTCTTCTTGGCAATCTTGGCAATGAGAGCCACA 330  
Qy 129 ArgValMetIleValGlyTyrIleCysAlaAlaIleAsnValAlaValPheAlaPro 148  
Db 391 CGGCTTAAGATTATCGAGAGATCTGTGCGATTCCTGTGTTGGTTCGCTGCTTANG 450  
Qy 149 LeuSerIleMetArgIleValIleIleTyrIleTyrIleTyrIleTyrIleTyrIleTyr 168  
Db 451 CTAAGCATATATCAGACGCGTAATTAAGACAGAGAGAGAGATGCTTACTTACTTA 510  
Qy 169 SerLeuPheLeuThrLeuCysAlaIleThrMetTyrPhePheTyrGlyPhePheTyrAsp 188  
Db 511 TCTTAACTTATCCATCACTGCTGCTCATATGCTCTTATATGCTTGTCTCAAGAC 570  
Qy 189 PheTyrIleAlaPheProAsnIleLeuGlyPheLeuPheGlyIleValGluMetLeuLeu 208  
Db 571 ATCTATGTTCTTCCGATGCTGCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 630  
Qy 209 TyrPheValTyrIleTyrIleTyrIleTyrIleTyrIleTyrIleTyrIleTyrIleTyr 220  
Db 631 TACGTTGCTTAAATACTGTAATAACGTCGCCATCTAGAGAGAGAGAGAGAGAGCT 690  
Qy 221 GluIleTyrIleTyrIleTyrIleTyrIleTyrIleTyrIleTyrIleTyrIleTyrIle 225  
Db 691 GCTTAAGTTACCGAGAGGAGCTGCTGATATGTTAGACTAGAGACAGATTATCCCTGAG 750  
Qy 226 -----ValArgGluAlaThrIleTyrSerIleGluGlyValGluIleIleAsn 241  
Db 751 CCAATCTCACTGTTGTTGTAAGGAGAACAGTGAATCTGCGGAATGATCGAAGGCTGAG 810  
Qy 242 IleGluAsp 244  
Db 811 ATTGAAGAT 819

Search completed: December 25, 2004, 01:08:07  
Job time : 5456 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 24, 2004, 19:39:56 ; Search time 135 Seconds  
(without alignments)  
1129.438 Million cell updates/sec

Title: US-09-743-885A-1  
Perfect score: 1353  
Sequence: 1 MAQLRADDSFTFGLLGNIV.....NSDNALQSMKDFSRRLRTSK 265

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Uniprot\_02:\*  
1: uniprot\_sprot:\*  
2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1353	100.0	265	2	Q9FPN0
2	704.5	52.0	258	2	Q9ZV02
3	703.5	52.0	258	2	Q8LDB6
4	656	48.5	268	2	P93332
5	647	47.8	319	2	Q6K602
6	623.5	46.1	292	2	Q8LAG0
7	622	46.0	292	2	Q9FY94
8	621.5	45.9	289	2	Q9SMN5
9	619.5	45.8	285	2	Q9LUB3
10	618	45.7	285	2	Q949G4
11	590	43.6	294	2	Q8LA22
12	590	43.6	294	2	Q9FG02
13	564.5	41.7	281	2	Q9SM25
14	539	39.8	307	2	Q61ZF3
15	539	39.8	307	2	BAD13102
16	539	39.8	307	2	BAD13102
17	462	34.1	241	2	Q84N33
18	462	34.1	241	2	Q84N33
19	447	33.0	247	2	Q8L907
20	440.5	32.6	273	2	Q8RZ08
21	435.5	32.2	230	2	Q9LUR4
22	435.5	32.2	230	2	AAQ65151
23	406	30.0	254	2	Q8W0K2
24	403.5	29.8	251	2	Q944M5
25	402.5	29.7	237	2	Q6L568
26	402.5	29.7	237	2	AA137996
27	401.5	29.7	259	2	Q8LR09
28	400.5	29.6	251	2	Q8LDB6
29	389	28.8	263	2	Q6N0M5
30	389	28.8	263	2	AAQ62417
31	388.5	28.7	259	2	Q6K4V2

32	383.5	28.3	258	2	Q8LBF7	Q8LBF7 arabidopsis
33	380	28.1	261	2	Q9FGL8	Q9FGL8 arabidopsis
34	370	27.3	238	2	Q70ET6	Q70ET6 lycopersico
35	370	27.3	238	2	CA647557	CA647557 lycopersi
36	363.5	26.9	239	2	Q8LPH5	Q8LPH5 arabidopsis
37	359.5	26.6	261	2	Q9C9M9	Q9C9M9 arabidopsis
38	355	26.2	176	2	Q6UA05	Q6UA05 gossypium b
39	355	26.2	176	2	AAQ84323	AAQ84323 gossypium
40	353.5	26.1	236	2	Q9LH79	Q9LH79 arabidopsis
41	350	25.9	202	2	Q9LPL1	Q9LPL1 arabidopsis
42	325	24.0	238	2	Q9SN64	Q9SN64 arabidopsis
43	321.5	23.8	213	2	Q9FM10	Q9FM10 arabidopsis
44	319.5	23.6	249	2	Q81628	Q81628 arabidopsis
45	313.5	23.2	263	2	Q9LRT5	Q9LRT5 arabidopsis

## ALIGNMENTS

RESULT 1					
ID	Q9FPN0	PRELIMINARY;	PRT;	265 AA.	
AC	Q9FPN0;				
DT	01-MAR-2001 (TrEMBLrel. 16, Created)				
DT	01-MAR-2001 (TrEMBLrel. 16, Last sequence update)				
DT	01-JUN-2003 (TrEMBLrel. 24, Last annotation update)				
DE	NEC1				
OS	Petunia hybrida (Petunia).				
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;				
OC	Lamiales; Solanales; Solanaceae; Petunia.				
NCBI	Taxid=4102;				
OX	[1]				
RN	SEQUENCE FROM N.A.				
RA	Ge Y.-X., Angenent G.C., Dahlhaus E., Franken J., Williams G.J.,				
RA	Creemers-Molenaar T.;				
RT	"Partial gene silencing of NEC1 results in early opening of anthers in				
RT	Petunia hybrida.";				
RL	Mol. Gen. Genet. 0:0-0(2001).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=20575716; PubMed=1135107;				
RA	Ge Y.-X., Angenent G.C., Wittlich P.E., Peter J., Franken J.,				
RA	Buscher M., Zhang L.-W., Dahlhaus E., Kater M.M., Williams G.J.,				
RA	Creemers-Molenaar T.;				
RT	"NEC1, a novel gene, highly expressed in nectary tissue of Petunia				
RT	hybrida.";				
RL	Plant J. 24:725-734(2000).				
DR	EMBL; AF13914; AAC34696.1; -;				
DR	GO; GO:0016020; C.membrane; IEA.				
DR	InterPro; IPR004316; MEN3_9lv.				
DR	Pfam; PF03083; MEN3_9lv; 2.				
DR	SEQUENCE 265 AA; 30256 MW; E6C6906B72A40C53 CRC64;				
Query Match 100.0%; Score 1353; DB 2; Length 265;					
Best Local Similarity Pred. No. 7.9e-99;					
Matches 265; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
QY	1	MAQLRADDSFTFGLLGNIVSEFWFLAPVPTFYKIKRKSSEGYQAI	PPVVALFSAGLL	60	
DB	1	MAQLRADDSFTFGLLGNIVSEFWFLAPVPTFYKIKRKSSEGYQAI	PPVVALFSAGLL	60	
QY	61	YVAVLRKNAVLYISNGFGCAITLTISLPLFAPRKSKITFTGMMLBLGALGMPT	120		
DB	61	YVAVLRKNAVLYISNGFGCAITLTISLPLFAPRKSKITFTGMMLBLGALGMPT	120		
QY	121	YLAAGSHRMVIMGICAAINVAFAAPLSIMRQVITKTSVEBPFTLSFLTLCTATMP	180		
DB	121	YLAAGSHRMVIMGICAAINVAFAAPLSIMRQVITKTSVEBPFTLSFLTLCTATMP	180		
QY	181	FYGFPPKDFYIAPPNITLGFGLGVQMLLVFVYKDSKRIDKSDPYREATKSGEVETII	240		
DB	181	FYGFPPKDFYIAPPNITLGFGLGVQMLLVFVYKDSKRIDKSDPYREATKSGEVETII	240		

QY	241	NIEDNSDNALQSMKDFSRRLRTSK	265
Db	241	NIEDNSDNALQSMKDFSRRLRTSK	265

  

RESULT 2			
ID	Q9ZV02	PRELIMINARY:	PRT; 258 AA.
AC	Q9ZV02:		
DT	01-MAY-1999	(TREMBLrel. 10, Created)	
DT	01-MAY-1999	(TREMBLrel. 10, Last sequence update)	
DT	01-JUN-2003	(TREMBLrel. 24, Last annotation update)	
DE	Similar to MCN3 protein.		
GN	Name=At2g39060;		
OS	Arabidopsis thaliana (Mouse-ear cress).		
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;		
OC	eurosid11; Brassicales; Brassicaceae; Arabidopsids.		
OX	NCBI_TaxID=3702;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RA	Rounsley S.D., Lin X., Kaul S., Shea T.P., Fujii C.Y., Mason T.M.,		
RA	Shen M., Roming C.M., Fraser C.M., Somerville C.R., Venter J.C.;		
RL	Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RA	Town C.D., Kaul S.;		
RL	Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; AC005770; AAC79616.1; -.		
DR	PIR; F84812; F84812.		
DR	GO; GO:0016020; C:membrane; IEA.		
DR	InterPro; IPR004316; MCN3_glv.		
DR	Pfam; PF03083; MCN3_glv_2.		
SQ	SEQUENCE 258 AA; 28716 MW; 0C58F140C586ADB CRC64;		

  

Query Match			
	52.1%;	Score 704.5;	DB 2; Length 258;
	Best Local Similarity	55.1%;	Pred. No.1.2e-47;
Matches	136;	Conservative	48; Mismatches 62; Indels 1; Gaps 1
QY	4	LRADLSTIFGLGNIVSFMFPLAVPPFFKTKYRKSSBEGVALPYMVALPSAGLLIYYA	63
Db	3	LKVEHIAELFGILGNIVSFGVFLSPVPFFYGGYKSSKGFOSIPIYICALMSATLLIYYG	62
QY	64	YLRKAYIVIVSINGCAIELTVYISLFTFYAPRKSKIFGWMLEL-ELGAGMWPITYL	122
Db	63	IMKTHAYIILISINTGCFETISYLFYLYLPAPREKAKITLKLIYVCTNIGLGLTLVNL	122
QY	123	LAEGSHRWIVGMICAINVAFAPLSIMROVITKTSVEFMPTLSFLTLCAATWFFY	182
Db	123	LVPKQHRVSTGVWCVAAYSLAVFASPLSVMRKVITKTSVEVMPFLISLTINAAVMFFY	182
QY	183	GFFKKDFPIAPNPNIIGLFGIVQMLLYVYKDSKIIDDEKSDPVAREATKSKGVEIINI	242
Db	183	GLLIKDKFIAPNPNIIGLFGVAQMILYMYQSTRTDLPTEINLANKTDVNEVPIVAVEL	242
QY	243	EDDNSDN 249	
Db	243	PDVGSN 249	

  

RESULT 3			
ID	Q8LDE6	PRELIMINARY:	PRT; 258 AA.
AC	Q8LDE6:		
DT	01-OCT-2002	(TREMBLrel. 22, Created)	
DT	01-OCT-2002	(TREMBLrel. 22, Last sequence update)	
DT	01-JUN-2003	(TREMBLrel. 24, Last annotation update)	
DE	Similar to MCN3 protein.		
OS	Arabidopsis thaliana (Mouse-ear cress).		
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;		
OC	eurosid11; Brassicales; Brassicaceae; Arabidopsids.		

[illegible]

SQ SEQUENCE 268 AA; 30081 MW; 3EBD39B2FAEF148E CRC64;  
 Query Match 48.5%; Score 656; DB 2; Length 268;  
 Best Local Similarity 48.5%; Pred. No. 8,4e-44;  
 Matches 129; Conservative 54; Mismatches 55; Indels 28; Gaps 6;  
 QY LSPFGLGIVNFMVFLAPVPTFYKIKRKSEGGYQAIIPYVVALFSAGLLYYAVALRK 68  
 DB LATFGILGNVISFLVFLAPVPTFYKIKRKSEGGYQAIIPYVVALFSAGLLYYAVALRK 67  
 QY AYLVISNGFCALIELTYISLFLFYAPRKSKIFTGMLM-LLELGALGMVPIITYLLAEGS 127  
 DB AFLITINSFGCVETIYIITLYIYAPRABNLTPKLSMANNQSPALLIIVTNVAVHGP 127  
 QY 128 HRVIMVGMICAINVAVPAAPLSIMROVITKSVENPFLLSLFTLCATMFFYGFKK 187  
 DB 128 LRVOVIGMVCVSLSVFAAPLSIVAVRTKSVFEPFNLSTLTLSATMFGYGFKK 187  
 QY 188 DFYIAPNIIIGLFGIVOMLIFYVYKD-SKRIDDEKSDVREATSKSKEGVETII----- 240  
 DB DITCIXLPNVIGXVGLGIQMLIYIRNGEKKAKKEKKAPI-BPPKS-----IVIEQLE 241  
 QY 241 -----NIEDNSDNALQSMEDFS 259  
 DB 242 KIEQEKKNKDDNE-----EKDKS 260  
 RESULT 5  
 ID Q6K602 PRELIMINARY; PRT; 319 AA.  
 AC Q6K602;  
 DT 05-JUL-2004 (TEMBLrel. 27, Created)  
 DT 05-JUL-2004 (TEMBLrel. 27, Last sequence update)  
 DE Putative nodulin 3.  
 GN Name=OJ1789 D08.23;  
 OS Oryza sativa (japonica cultivar-group).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Eriharctideae; Oryzaceae; Oryza.  
 NC NCBI\_TaxID=39947;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Sasaki T., Matsumoto T., Katayose Y.;  
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AP005299; BAD23245.1; -.  
 DR InterPro; IPR004316; MCN3\_s1v.  
 DR Pfam; PF03083; MCN3\_s1v; 2.  
 SQ SEQUENCE 319 AA; 34964 MW; 308075F4787795CA CRC64;  
 Query Match 47.8%; Score 647; DB 2; Length 319;  
 Best Local Similarity 50.4%; Pred. No. 5.1e-43;  
 Matches 116; Conservative 53; Mismatches 55; Indels 6; Gaps 2;  
 QY 10 SFIFGLIGNIVSFMVFLAPVPTFYKIKRKSEGGYQAIIPYVVALFSAGLLYYAVALRK 69  
 DB AFIFGLIGNISLWVFLAPVPTFYKIKRKSEGGYQAIIPYVVALFSAGLLYYAVALRK 71  
 QY 70 YLVISNGFCALIELTYISLFLFYAPRKSKIFTGMLM-LLELGALGMVPIITYLLAEGS 128  
 DB ELVITINSFGCVETIYIITLYIYAPRABNLTPKLSMANNQSPALLIIVTNVAVHGP 127  
 QY 129 RVMIVGMICAINVAVPAAPLSIMROVITKSVENPFLLSLFTLCATMFFYGFKK 188  
 DB 132 RVHVLGMCVAVSVFAAPLSIRLVIRTKSVFEPFNLSTLTLSATMFGYGFKK 191  
 QY 189 FYIAPNIIIGLFGIVOMLIFYVYKDSKRIDDEKSDVREATSKSKEGVETII----- 238  
 DB 192 VFVALPVLGVFVGAQALMAYRSKKPLVASSSAVVAH-----GUEI 236  
 RESULT 6  
 ID Q8LAGO  
 AC Q8LAGO;  
 DT 01-OCT-2002 (TEMBLrel. 22, Created)  
 DT 01-OCT-2002 (TEMBLrel. 22, Last sequence update)  
 DE Senescence-associated protein (SAG29).  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosid II; Brassicales; Brassicaceae; Arabidopsis.  
 NC NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=22088475; PubMed=12093376;  
 RX Haas B.J., Volkovskiy N., Town C.D., Troukhan M., Alexandrov N.,  
 RA Feldmann K.A., Flavell R.B., White O., Salzman S.L.;  
 RT "Full-length messenger RNA sequences greatly improve genome  
 annotation."  
 RL Genome Biol. 3:RESEARCH0029-RESEARCH0029(2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,  
 RA Feldmann K.;  
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AY087836; AAM65389.1; -.  
 DR GO; GO:0016020; C-membrane; IEA.  
 DR InterPro; IPR002114; HPR\_Serp\_S.  
 DR InterPro; IPR004316; MCN3\_s1v.  
 DR Pfam; PF03083; MCN3\_s1v; 2.  
 DR PROSITE; PS00589; PTS\_HPR\_SER; UNKNOWN 1.  
 SQ SEQUENCE 292 AA; 32919 MW; 4D92FCA710412P55 CRC64;  
 Query Match 46.1%; Score 623.5; DB 2; Length 292;  
 Best Local Similarity 43.9%; Pred. No. 3.4e-41;  
 Matches 119; Conservative 58; Mismatches 63; Indels 31; Gaps 4;  
 QY 9 LSPFGLGIVNFMVFLAPVPTFYKIKRKSEGGYQAIIPYVVALFSAGLLYYAVALRK 68  
 DB LATFGILGNVISFLVFLAPVPTFYKIKRKSEGGYQAIIPYVVALFSAGLLYYAVALRK 69  
 QY AYLVISNGFCALIELTYISLFLFYAPRKSKIFTGMLM-LLELGALGMVPIITYLLAEGS 127  
 DB AFLITINSFGCVETIYIITLYIYAPRABNLTPKLSMANNQSPALLIIVTNVAVHGP 127  
 QY 128 HRVIMVGMICAINVAVPAAPLSIMROVITKSVENPFLLSLFTLCATMFFYGFKK 186  
 DB 130 PLOVSVLGMICVAVSVFAAPLSIVAVRTKSVFEPFNLSTLTLSATMFGYGFKK 189  
 QY 187 KDFYIAPNIIIGLFGIVOMLIFYVYKDSKR----- 218  
 DB 190 NDCIAPNVGVGLQVLLVLRNNSNEKINSSEBOLKSTIVMSPLGVSEVHV 249  
 QY 219 DDEKSDVREATSKSKEGVETIINIEDNSDN 249  
 DB 250 VTESVDPLSEAV-HHEDLSKVTKEEPSIEN 279  
 RESULT 7  
 ID Q9FY94  
 AC Q9FY94;  
 DT 01-MAR-2001 (TEMBLrel. 16, Created)  
 DT 01-MAR-2001 (TEMBLrel. 16, Last sequence update)  
 DE Putative senescence-associated protein SAG29 (SAG29).  
 GN Name=119L5.130; Synonyms=At5g13170;  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosid II; Brassicales; Brassicaceae; Arabidopsis.  
 NC NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.

ID Q8LAGO PRELIMINARY; PRT; 292 AA.  
 AC Q8LAGO;  
 DT 01-OCT-2002 (TEMBLrel. 22, Created)  
 DT 01-OCT-2002 (TEMBLrel. 22, Last sequence update)  
 DE Senescence-associated protein (SAG29).  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosid II; Brassicales; Brassicaceae; Arabidopsis.  
 NC NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=22088475; PubMed=12093376;  
 RX Haas B.J., Volkovskiy N., Town C.D., Troukhan M., Alexandrov N.,  
 RA Feldmann K.A., Flavell R.B., White O., Salzman S.L.;  
 RT "Full-length messenger RNA sequences greatly improve genome  
 annotation."  
 RL Genome Biol. 3:RESEARCH0029-RESEARCH0029(2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,  
 RA Feldmann K.;  
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AY087836; AAM65389.1; -.  
 DR GO; GO:0016020; C-membrane; IEA.  
 DR InterPro; IPR002114; HPR\_Serp\_S.  
 DR InterPro; IPR004316; MCN3\_s1v.  
 DR Pfam; PF03083; MCN3\_s1v; 2.  
 DR PROSITE; PS00589; PTS\_HPR\_SER; UNKNOWN 1.  
 SQ SEQUENCE 292 AA; 32919 MW; 4D92FCA710412P55 CRC64;  
 Query Match 46.1%; Score 623.5; DB 2; Length 292;  
 Best Local Similarity 43.9%; Pred. No. 3.4e-41;  
 Matches 119; Conservative 58; Mismatches 63; Indels 31; Gaps 4;  
 QY 9 LSPFGLGIVNFMVFLAPVPTFYKIKRKSEGGYQAIIPYVVALFSAGLLYYAVALRK 68  
 DB LATFGILGNVISFLVFLAPVPTFYKIKRKSEGGYQAIIPYVVALFSAGLLYYAVALRK 69  
 QY AYLVISNGFCALIELTYISLFLFYAPRKSKIFTGMLM-LLELGALGMVPIITYLLAEGS 127  
 DB AFLITINSFGCVETIYIITLYIYAPRABNLTPKLSMANNQSPALLIIVTNVAVHGP 127  
 QY 128 HRVIMVGMICAINVAVPAAPLSIMROVITKSVENPFLLSLFTLCATMFFYGFKK 186  
 DB 130 PLOVSVLGMICVAVSVFAAPLSIVAVRTKSVFEPFNLSTLTLSATMFGYGFKK 189  
 QY 187 KDFYIAPNIIIGLFGIVOMLIFYVYKDSKR----- 218  
 DB 190 NDCIAPNVGVGLQVLLVLRNNSNEKINSSEBOLKSTIVMSPLGVSEVHV 249  
 QY 219 DDEKSDVREATSKSKEGVETIINIEDNSDN 249  
 DB 250 VTESVDPLSEAV-HHEDLSKVTKEEPSIEN 279  
 RESULT 7  
 ID Q9FY94  
 AC Q9FY94;  
 DT 01-MAR-2001 (TEMBLrel. 16, Created)  
 DT 01-MAR-2001 (TEMBLrel. 16, Last sequence update)  
 DE Putative senescence-associated protein SAG29 (SAG29).  
 GN Name=119L5.130; Synonyms=At5g13170;  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosid II; Brassicales; Brassicaceae; Arabidopsis.  
 NC NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RA Bevan M., Rose M., Hempel S., Entian K.-D., Bancroft I., Mewes H.W.,  
RA Rudd S., Lemcke K., Mayer K.F.X.;  
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
RN [12]  
RP SEQUENCE FROM N.A.  
RA EU Arabidopsis sequencing project;  
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
RN [13]  
RP SEQUENCE FROM N.A.  
RA Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Chung M.K.,  
RA Goldsmith A.D., Lee J.M., Quach H.L., Toriumi M., Yu G., Bowser L.,  
RA Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,  
RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,  
RA Miranda M., Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M.,  
RA Seki M., Shinn P., Southwick A., Shinozaki K., Davis R.W., Ecker J.R.,  
RA Theologis A.;  
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
RN [14]  
RP SEQUENCE FROM N.A.  
RA Yamada K., Banh J., Chan M.M., Chang C.H., Chang E., Dale J.M.,  
RA Deng J.M., Goldsmith A.D., Lee J.M., Onodera C.S., Quach H.L.,  
RA Tang C., Toriumi M., Wu H.C., Yamamura Y., Yu G., Bowser L.,  
RA Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,  
RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,  
RA Miranda M., Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M.,  
RA Seki M., Shinn P., Southwick A., Shinozaki K., Davis R.W., Ecker J.R.,  
RA Theologis A.;  
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.  
RN [15]  
RP SEQUENCE FROM N.A.  
RA EMBL; AL313711; CAC05445.1; -  
DR EMBL; AY045949; AAK76623.1; -  
DR EMBL; AY113934; AAM44982.1; -  
DR GO; GO:0016020; C:membrane; IEA.  
DR InterPro; IPR002114; HPR Serp S.  
DR InterPro; IPR004316; MCN3 sly-2.  
DR Pfam; PF03083; MCN3 sly-2.  
DR PROSITE; PS00589; PTS\_HPR\_SER; UNKNOWN 1.  
SQ SEQUENCE 292 AA; 32936 MW; 4C3826F6D3A845D5 CRC64;  
Query Match 46.1%; Score 623.5; DB 2; Length 292;  
Best Local Similarity 43.9%; Pred. No. 3,4e-41;  
Matches 119; Conservative 58; Mismatches 63; Indels 31; Gaps 4;  
QY 9 LSPFGLGNGVSPVFLAPPTFYKTKYKSSSEGYQAIIPVVALFSAAGLLIYAYIRKN 68  
DB 10 LAFIFGLIGNVISPVLAPVPTFYRIYKRKSTESFQSLPYQVSLFSCMLWLYALIKD 69  
QY 69 AYLIVSIFGCGALBELTYISLFLFAPRKSIFPGMLL-LELQALGVMPTLTLAEGS 127  
DB 70 AFLIITINSFCCVVELYIAIFFAVATREKRISAMKLFIAINVVAFPSLILWVTHFVAKTP 129  
QY 128 H-RVMIYVGMICAAINAVFAAPLISIMROVITKTSVEFMPFTLSFLTLCATMWFYGF 186  
DB 130 PLQSVLGMICVAVLSVFAAPLMIIVARVITKTSVEFMPFTLSFLTISATMWAAYGFL 189  
QY 187 KDFIAPFNITIGFLPGIVQMLLYVYKDSK-----I 218  
DB 190 NDICIAIPNVVGFYGLQWLDLYVYRNSNEKPEKINSSEQQLKSIVMSPLGVSEVHPV 249  
QY 219 DDEKSDPYREATKSEGEVETIINIEDNSN 249  
DB 250 VTSEVDPLSEAV-KHEDLSKVTKEEPEIEN 279

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Choisme N., Robert C., Brottier P., Wincker P., Catolico L., K.,  
RA Artiguenave F., Saurin W., Weissendach J., Mewes H.W., Lemcke K.,  
RA Mayer K.F.X., Quetier F., Salanoubat M.;  
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.  
RN [12]  
RP SEQUENCE FROM N.A.  
RA EU Arabidopsis sequencing project;  
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.  
RN [13]  
RP SEQUENCE FROM N.A.  
RA Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Chung M.K.,  
RA Goldsmith A.D., Lee J.M., Quach H.L., Toriumi M., Yu G., Bowser L.,  
RA Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,  
RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,  
RA Miranda M., Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M.,  
RA Seki M., Shinn P., Southwick A., Shinozaki K., Davis R.W., Ecker J.R.,  
RA Theologis A.;  
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.  
RN [14]  
RP SEQUENCE FROM N.A.  
RA Cheuk R., Chen H., Kim C.J., Meyers M.C., Banh J., Bowser L.,  
RA Carninci P., Chang E., Dale J.M., Goldsmith A.D., Hayashizaki Y.,  
RA Ishida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Lam B.,  
RA Lee J.M., Lin J., Miranda M., Narusaka M., Nguyen M., Onodera C.S.,  
RA Palm C.J., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,  
RA Tang C.C., Toriumi M., Wu H.C., Yamada K., Yamamura Y., Yu G., Yu S.,  
RA Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;  
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.  
RN [15]  
RP SEQUENCE FROM N.A.  
RA Yamada K., Banh J., Chan M.M., Chang C.H., Chang E., Dale J.M.,  
RA Deng J.M., Goldsmith A.D., Lee J.M., Onodera C.S., Quach H.L.,  
RA Tang C., Toriumi M., Wu H.C., Yamamura Y., Yu G., Bowser L.,  
RA Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,  
RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,  
RA Meyers M.C., Miranda M., Narusaka M., Nguyen M., Palm C.J.,  
RA Sakurai T., Satou M., Seki M., Shinn P., Southwick A., Shinozaki K.,  
RA Davis R.W., Ecker J.R., Theologis A.;  
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
RN [16]  
RP SEQUENCE FROM N.A.  
RA Cheuk R., Chen H., Kim C.J., Shinn P., Banh J., Bowser L.,  
RA Carninci P., Chang E., Dale J.M., Goldsmith A.D., Hayashizaki Y., Ishida J.,  
RA Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Lam B., Lee J.M.,  
RA Lin J., Liu S.X., Miranda M., Narusaka M., Nguyen M., Palm C.J.,  
RA Pham P.K., Quach H.L., Sakano H., Sakurai T., Satou M., Seki M.,  
RA Southwick A., Toriumi M., Yamada K., Yu G., Shinozaki K., Davis R.W.,  
RA Theologis A., Ecker J.R.;  
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.  
RN [17]  
RP SEQUENCE FROM N.A.  
RA Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Banh J.,  
RA Bowser L., Carninci P., Dale J.M., Goldsmith A.D., Hayashizaki Y.,  
RA Ishida J., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G.,  
RA Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,  
RA Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T.,  
RA Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,  
RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,  
RA Ecker J.R.;  
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
RN [18]  
RP SEQUENCE FROM N.A.  
DR EMBL; AL13315; CAB62363.1; -  
DR EMBL; AY070412; AAL49908.1; -  
DR EMBL; AY078041; AAL77742.1; -  
DR EMBL; AY096594; AAM20244.1; -  
DR EMBL; AF361825; AAK32837.1; -  
DR EMBL; AF419559; AAL31891.1; -  
DR PIR; T46218; T46218.



QY 1 MAQIRADSLPIFGILGNIVSFVWFLAPPTFYKTKYKSSGQYQAIPTVVALFSAQLL 60  
 DB 1 MAMTR-ESMAFVFGIINIIISFAVFLSPPLTFYVIFPKKSAEGQALPYVALFSAQLLMI 59  
 QY 61 YVAVL-RKNAVLYISINGPGCAIELEYISLFLFAPARKSKIFT-GWMLLELGLGVMMP 118  
 DB 60 YVAVKESKSLTLTITPTFGVESAATIMEFLYAPKQKSTIKLLILNFGFGMLL 119  
 QY 119 ITVLLAGSHRWIVGICAINVAFAPLSIRQVITKTSVEFMFTLSLFLTCATM 178  
 DB 120 STLVLSGAKRLAIIGMICLVNISVFAAPLFAVISKIRSVEMPFISFLTINAVM 179  
 QY 179 WFFFGFKKPFYIAFPNITLFGIYOMLYFYVYKSKRIDDEKSDVREATSKGEVEI 238  
 DB 180 WFFYGLLRDYVALPPTLGFVFGIIQMVVLYLRN-----ATPEVKA--FMKQEL 229  
 QY 239 -----INIEDNSDN 249  
 DB 230 SGCHIDVVKIGTDS 244

## RESULT 11

ID 082587 PRELIMINARY; PRT; 285 AA.

AC 082587; (TREMBlrel. 08, Created)

DT 01-NOV-1998 (TREMBlrel. 08, last sequence update)

DT 05-JUN-2004 (TREMBlrel. 27, last annotation update)

DE Pucative Mtn3 protein.

GN Name=Mtn3; Synonyms=At5g23660;

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;

OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.

OX NCBI\_Taxid=3702;

RP SEQUENCE FROM N.A.

RA Cheong J.-J.;

RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RA Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Chung M.K.,

RA Goldsmith A.D., Lee J.M., Quach H.L., Toriumi M., Yu G., Bowser L.,

RA Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,

RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,

RA Miranda M., Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M.,

RA Seki M., Shim P., Southwick A., Shinozaki K., Davis R.W., Ecker J.R.,

RA Theologis A.;

RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.

RN [4]

RP SEQUENCE FROM N.A.

RA Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Etgu P., Lee J.M.,

RA Toriumi M., Yu G., Brooks S., Chao Q., Chen H., Karlin-Neumann G.,

RA Kim C., Lam B., Miranda M., Nguyen M., Palm C.J., Shim P.,

RA Southwick A., Davis R.W., Ecker J.R., Theologis A.;

RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.

RN [5]

RP SEQUENCE FROM N.A.

RA Yamada K., Banh J., Chan M.M., Chang C.H., Chang E., Dale J.M.,

RA Deng J.M., Goldsmith A.D., Lee J.M., Onodera C.S., Quach H.L.,

RA Tang C., Toriumi M., Wu H.C., Yamamura Y., Yu G., Bowser L.,

RA Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,

RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,

RA Meyers W.C., Miranda M., Narusaka M., Nguyen M., Palm C.J.,

RA Sakurai T., Satou M., Seki M., Shim P., Southwick A., Shinozaki K.,  
 RA Davis R.W., Ecker J.R., Theologis A.;

RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF095641; AAC64192.1; -

DR EMBL; AB025633; BAA69723.1; -

DR EMBL; AY057575; AAL09814.1; -

DR EMBL; AY059108; AAL15214.1; -

DR EMBL; AY116672; AAM47150.1; -

DR PIR; T51837; T51837.

DR GO; GO:0016020; C:membrane; IEA.

DR InterPro; IPR004316; Mtn3\_slv.

DR Pfam; PF03083; Mtn3\_slv; 2

SQ SEQUENCE 285 AA; 31486 MW; E18CD08B1D247E03 CRC64;

Query Match 45.7%; Score 618; DB 2; Length 285;

Best Local Similarity 47.1%; Pred. No. 8.9e-41;

Matches 123; Conservative 54; Mismatches 62; Indels 22; Gaps 4;

QY 10 SPFGILGNIVSFVWFLAPPTFYKTKYKSSGQYQAIPTVVALFSAQLLTYAFLRNA 69

DB 11 AFVFGILGNLISFAVFLSPPLTFYVIFPKKSAEGQALPYVALFSAQLLMIYATOKDV 70

QY 70 YLVSINGPGCAIELEYISLFLFAPARKSKIFT-GWMLLELGLGVMMPITYLLAGSH 128

DB 71 FLAVTINSFCFLETITYISFVAFASKKAMLTVKLLLNFGGFCILILCOPLAKGT 130

QY 129 RWIVGICAINVAFAAPLSIRQVITKTSVEFMFTLSLFLTCATMWFYGFKKD 188

DB 131 RAKIIGICVGFVCFAPAPLSIRQVITKTSVEFMFTLSLFLTCATMWFYGFKKD 190

QY 189 FYIAPNITLFGIYOMLYFYVYKSKRIDDEKSDVREATSKGEVEI-KSKE 234

DB 191 IYVAFVIVGIFVGAQMLLYVYKCKTPPSDLVEKELEAKLPEVSDVWKLGTLSPE 250

QY 235 GVEI-----NIEDNSD 248

DB 251 PVATIVRSVNTCNCNDRAE 271

## RESULT 12

ID 08LAZ2 PRELIMINARY; PRT; 294 AA.

AC 08LAZ2; (TREMBlrel. 22, Created)

DT 01-OCT-2002 (TREMBlrel. 22, last sequence update)

DT 01-JUN-2003 (TREMBlrel. 24, last annotation update)

DE Mtn3-like protein.

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;

OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.

OX NCBI\_Taxid=3702;

RP SEQUENCE FROM N.A.

RA Haas B.U., Voltofsky N., Town C.D., Troukhan M., Alexandrov N.,

RA Feldmann K.A., Flavell R.B., White O., Salzberg S.L.;

RL "Full-length messenger RNA sequences greatly improve genome

RT annotation".

RL Genome Biol. 3:RESEARCH0029-RESEARCH0029(2002).

RN [2]

RP SEQUENCE FROM N.A.

RA Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,

RA Feldmann K.;

RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; AY087516; AAM65058.1; -

DR GO; GO:0016020; C:membrane; IEA.

DR InterPro; IPR004316; Mtn3\_slv.

DR Pfam; PF03083; Mtn3\_slv; 2.

SQ SEQUENCE 294 AA; 32503 MW; 64794865A0F0B84 CRC64;

Query Match 43.6%; Score 590; DB 2; Length 294;

Best Local Similarity 48.6%; Pred. No. 1.5e-38;



Matches 126; Conservative 46; Mismatches 57; Indels 30; Gaps 6;

QY 10 SFIFGLGNIVSMVFLAPVPTFYKTKYKRSSEGYQAIPYVVALFSAGLLYYAVALRK-N 68  
 DB 9 AFVFGILGNISFVFLAPVPTFVRICKKSTEGFSLPVSALFSAIMLYIAMQDGT 68  
 QY 69 AYLIVSINGFGCAIBETIYISLFLFYAPRKSIFT---GWLMLLELGALGMVPIITYLLA 124  
 DB 69 AFLITITINAGVCIEITIIYIVLFSVYANKKTRISTLKVGLNLFGLFAIIVLCE---LIT 125  
 QY 125 EGSHRVIMVGMICAINAVFAAPLISIMROVITKTSVEFMPFTLSLFLTCATMFFYGF 184  
 DB 126 KSGTREKVLGICVGSVSFAFPLSIRVVRTRSEFMPFSLFLITISAVTWLIFYG 185  
 QY 185 FKDFEIAFPNITIGLFGIVQMLLYFYVYKSDKRIIDDEKSPVBEATKSKGEVITINIED 244  
 DB 186 AIKDFYVALPNNVGAFLGAVQMLIYIIFKYYKTPVAQKTD-----KSK----- 228  
 QY 245 DNSDNALQSMKDFSRILT 263  
 DB 229 DVSDBSI-----DIKLT 242

## RESULT 13

Q9FGQ2 PRELIMINARY; PRT; 294 AA.

AC 09FGQ2;  
 DT 01-MAR-2001 (TEMBLrel. 16, Created)  
 DT 01-MAR-2001 (TEMBLrel. 16, Last sequence update)  
 DT 01-JUN-2003 (TEMBLrel. 24, Last annotation update)  
 DE MCN3-like protein.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosida II; Brassicales; Brassicaceae; Arabidopsids.  
 NCBI\_TaxId=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Kaneko T., Katoh T., Asamizu E., Sato S., Nakamura Y., Kotani H.,  
 RA Tabata S.;  
 RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AB025617; BAB08903.1; -  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR InterPro; IPR004316; MCN3\_slv.  
 DR Pfam; PF03083; MCN3\_slv; 2.  
 SQ SEQUENCE 294 AA; 32503 MW; 26D88F65398D0DAE CRC64;

Query Match 43.6%; Score 590; DB 2; Length 294;  
 Best Local Similarity 48.6%; Pred. No. 1.5e-38;  
 Matches 126; Conservative 46; Mismatches 57; Indels 30; Gaps 6;

QY 10 SFIFGLGNIVSMVFLAPVPTFYKTKYKRSSEGYQAIPYVVALFSAGLLYYAVALRK-N 68  
 DB 9 AFVFGILGNISFVFLAPVPTFVRICKKSTEGFSLPVSALFSAIMLYIAMQDGT 68  
 QY 69 AYLIVSINGFGCAIBETIYISLFLFYAPRKSIFT---GWLMLLELGALGMVPIITYLLA 124  
 DB 69 AFLITITINAGVCIEITIIYIVLFSVYANKKTRISTLKVGLNLFGLFAIIVLCE---LIT 125  
 QY 125 EGSHRVIMVGMICAINAVFAAPLISIMROVITKTSVEFMPFTLSLFLTCATMFFYGF 184  
 DB 126 KSGTREKVLGICVGSVSFAFPLSIRVVRTRSEFMPFSLFLITISAVTWLIFYG 185  
 QY 185 FKDFEIAFPNITIGLFGIVQMLLYFYVYKSDKRIIDDEKSPVBEATKSKGEVITINIED 244  
 DB 186 AIKDFYVALPNNVGAFLGAVQMLIYIIFKYYKTPVAQKTD-----KSK----- 228  
 QY 245 DNSDNALQSMKDFSRILT 263  
 DB 229 DVSDBSI-----DIKLT 242

## RESULT 14

Q9SM25 PRELIMINARY; PRT; 281 AA.

AC 09SM25;  
 DT 01-MAY-2000 (TEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TEMBLrel. 13, Last sequence update)  
 DT 05-JUL-2004 (TEMBLrel. 27, Last annotation update)  
 DE MCN3-like protein.  
 GN Name=F13M23.150; Synonyms=AT4g25010;  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosida II; Brassicales; Brassicaceae; Arabidopsids.  
 NCBI\_TaxId=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Bevan M., Medler H., Medler E., Wambutt R., Hoheisel J., Mewes H.W.,  
 RA Mayer K.F.X., Scheller C.;  
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Medler H., Medler E., Wambutt R., Mewes H.W., Lemcke K., Mayer K.F.X.;  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RA EU Arabidopsis sequencing project;  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AL035523; CAB36743.1; -  
 DR EMBL; AL161562; CAB79410.1; -  
 DR PIR; T05522; T05522.  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR InterPro; IPR004316; MCN3\_slv.  
 DR Pfam; PF03083; MCN3\_slv; 2.  
 SQ SEQUENCE 281 AA; 30969 MW; D01DA2649AC9B92 CRC64;

Query Match 41.7%; Score 564.5; DB 2; Length 281;  
 Best Local Similarity 50.8%; Pred. No. 1.5e-36;  
 Matches 121; Conservative 39; Mismatches 65; Indels 13; Gaps 5;

QY 9 LSIFGLGNIVSMVFLAPVPTFYKTKYKRSSEGYQAIPYVVALFSAGLLYYAVALRK-N 68  
 DB 8 LAVTREVGNISFVFLAPVPTFVRICKKSTEGFSLPVSALFSAIMLYIAMQDGT 67  
 QY 69 A-YLIVSINGFGCAIBETIYISLFLFYAPRKSIFT---GWLMLLELGALGMVPIITYLL 123  
 DB 68 AFLITITINAGVCIEITIIYIVLFSVYANKKTRISTLKVGLNLFGLFAIIVLCE---LIT 124  
 QY 124 AEGSHRVIMVGMICAINAVFAAPLISIMROVITKTSVEFMPFTLSLFLTCATMFFYGF 183  
 DB 125 TKSNSEKVLGICVGSVSFAFPLSIRVVRTRSEFMPFSLFLITISAVTWLIFYG 184  
 QY 184 FKDFEIAFPNITIGLFGIVQMLLYFYVYKSDKRIIDDEKSPVBEATKSKGEVITINIED 239  
 DB 185 LAIKDFYVALPNNVGAFLGAVQMLIYIIFKYYKTPVAQKTD-----KIVSDHSIMNV 239

## RESULT 15

Q6YZF3 PRELIMINARY; PRT; 307 AA.

AC 06YZF3;  
 DT 05-JUL-2004 (TEMBLrel. 27, Created)  
 DT 05-JUL-2004 (TEMBLrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TEMBLrel. 27, Last annotation update)  
 DE Putative MCN3.  
 GN Name=P0702C09.33; Synonyms=OSJNBa0033D24.17;  
 OS Oryza sativa (japonica cultivar-group).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzaceae; Oryza.  
 NCBI\_TaxId=39947;

RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Sasaki T., Matsumoto T., Katayose Y.;  
 RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Sasaki T., Matsumoto T., Katayose Y.;  
 RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AP005528; BAD13168.1; -;  
 DR EMBL; AP005439; BAD13102.1; -;  
 DR InterPro: IPR004316; MEN3\_glv.  
 DR Pfam; PF03083; MEN3\_glv; 2.  
 SQ SEQUENCE 307 AA; 32978 MW; 6E505D8DA701B664 CRC64;

Query Match 39.8%; Score 539; DB 2; Length 307;  
 Best Local Similarity 45.4%; Pred. No. 1.6e-34;  
 Matches 104; Conservative 49; Mismatches 70; Indels 6; Gaps 2;

QY 14 GLUGNIVSPNVFLAPPTFYKIYKRKSEGYQALPYMVALFSAGLLLYAYLRNAYLIV 73  
 Db 17 GVAIGNIISFLVFLAPVATFLQYVKKSTGGYSSVPYVVALFSSVLWIFVALVKTNRPIL 76  
 QY 74 SINGFGCAIETYSLSLFLYAPRKSIFTCMLMLELGLGVMPTIYLLAEGSHRMI 132  
 Db 77 TINAFGCGVEAAYIVLVYAPRRARLTALFLLLDVAAPALIVTTLYIVPRHQVKF 136  
 QY 133 VGMICAIINAVFAAPLISIMROVITKTSVEFMPTLSLFTLCATMWFYGFYKDFYIA 192  
 Db 137 LGSVCLAFSNAVFAPISTIIFKVIKTSVEFMPIGLSVCLTISAVAMFCYGLFTKDPYV 196  
 QY 193 FPNILGFLFGIVQMLIFVYKDSKR----IDDEKSDPVREATKSKEGV 236  
 Db 197 YPNVGFFFCVQMGLEYFYWRKPRNTAVLPPTSDMSPISSAAAAATQRV 245

Search completed: December 24, 2004, 20:43:08  
 Job time : 138 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using bw model

Run on: December 24, 2004, 20:29:56 ; Search time 52 Seconds  
(Without alignments)  
490.335 Million cell updates/sec

Title: US-09-743-885A-1

Perfect score: 1353  
Sequence: 1 MAQLRADDSFTFGLGNIV.....NSDNALQSMKQDSRLRTSK 265

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR 79: \*  
2: PIR 79: \*  
3: PIR 79: \*  
4: PIR 79: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	704.5	52.1	258	2	F84812	similar to MEN3 pr
2	622	46.0	289	2	T46218	MTN3-like protein
3	618	45.7	285	2	T51837	MTN3 homolog [impo
4	564.5	41.7	281	2	T05522	hypothetical prote
5	350	25.9	202	2	F86347	hypothetical prote
6	325	24.0	238	2	T04280	hypothetical prote
7	319.5	23.6	249	2	T01891	hypothetical prote
8	291.5	21.5	238	2	A71425	hypothetical prote
9	243	18.0	355	2	T19016	hypothetical prote
10	233.5	17.3	686	2	A71607	MEN3/RAG1P-like p
11	220	16.3	250	2	T26725	hypothetical prote
12	216	16.0	221	2	UC4761	recombinational acti
13	213.5	15.8	233	2	T31865	hypothetical prote
14	209.5	15.5	299	2	T32982	hypothetical prote
15	146	10.8	224	2	T32655	hypothetical prote
16	114.5	8.5	300	2	S36430	hypothetical prote
17	114	8.4	546	1	B24707	NADH2 dehydrogenas
18	112	8.3	546	2	A69880	hypothetical prote
19	109	8.1	808	2	T16564	hypothetical prote
20	108.5	8.0	567	2	B81408	probable thiol-dis
21	108.5	8.0	687	2	C71849	carbon starvation
22	108	8.0	583	2	T48473	amino acid transpo
23	107.5	7.9	687	2	H64655	carbon starvation
24	106	7.8	475	2	B29606	methylesterase A r
25	105.5	7.8	361	2	D89869	conserved hypochet
26	104.5	7.7	228	2	D90513	hypothetical prote
27	103.5	7.6	489	2	G90402	hypothetical prote
28	102	7.5	253	2	E75055	hypothetical prote
29	100.5	7.4	229	2	T24122	hypothetical prote

30	100.5	7.4	356	2	B97822	probable permease
31	100	7.4	444	2	T12297	NADH2 dehydrogenas
32	99.5	7.4	399	2	AF3249	hypothetical prote
33	98.5	7.3	395	2	H81390	probable integral
34	98.5	7.3	567	2	D65043	hypothetical prote
35	98.5	7.3	567	2	T08653	hypothetical prote
36	98.5	7.3	829	2	A93960	hypothetical prote
37	98	7.2	202	2	F72008	CT852 hypothetical
38	97.5	7.2	201	2	H97218	uncharacterized co
39	97.5	7.2	508	2	D90567	hypothetical prote
40	97.5	7.2	596	2	S46001	probable amino aci
41	97	7.2	245	2	C30010	hypothetical ORF-6
42	97	7.2	380	2	D64352	hypothetical prote
43	97	7.2	583	2	AH1151	glycerophosphoryl
44	96.5	7.1	422	2	AC1189	phosphotransferase
45	96.5	7.1	433	2	T11162	NADH2 dehydrogenas

#### ALIGNMENTS

RESULT 1  
F84812  
similar to MEN3 protein [imported] - Arabidopsis thaliana  
C/Species: Arabidopsis thaliana (mouse-ear creas)  
C/Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 09-Jul-2004  
C/Accession: F84812  
R/Lin, X.; Kaul, S.; Rounsley, S.D.; Shua, T.P.; Benito, M.T.; Town, C.D.; Fujii, C.Y.; M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.; Euse, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J. Nature 402, 761-768, 1999  
A/Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
A/Reference number: A84420; MID:20083487; PMID:1061197  
A/Accession: F84812  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-258 <STO>  
A/Cross-references: UNIPROT:Q92V02; GB:AE002093; NID:G3928090; PIDN:AA079616.1; GSPDB:GNK  
C/Genes: A;Gene: At2g39060  
A/Map position: 2

Query Match 52.1%; Score 704.5; DB 2; Length 258;  
Best Local Similarity 55.1%; Pred. No. 3.8e-55;

Matches 136; Conservative 48; Mismatches 62; Indels 1; Gaps 1;

Qy	4	LRADDSTFIFGLGNIVSFVFLAPVPTFYKIKRKSSEGCQAIPIYVVALFSAGLLLYA	63
Db	3	LKVHEIAFLFGLGNIVSFGVFLSPVPTFYGIKKSSKGFQIPICALASATLLLYG	62
Qy	64	YLRKNAVLYVINGFCATFETYSFLFYAPPKSKITFGWMLL-ELGALGVMPITTY	122
Db	63	IMKTHAVLIISITFCFETISYLFYIYAPREAKISTLKLIVINIGLGLIILVNL	122
Qy	123	LAGSHRWMTVGWICAINAVFAAPLSTIRGVKTSVEFPPTSLFTLCATMFFY	182
Db	123	LTVKQHRVSTGVWCAVSLAVFASPLSVWRKVIKKSVEYMPFLSLSTITLNAVMEFY	182
Qy	183	GFFKDFYIAFPNIIPLFQIVQMLLYFYVYKDSKRIIDEDKSDPVREATSKSGEVEIINI	242
Db	183	GLIKDKFTAMPILDFLFGVAQMILYMYVQGSTKTDLPFENQLANKTIDVNEPIYAVEL	242
Qy	243	EDDNDN 249	
Db	243	PDVGSND 249	

RESULT 2  
T46218  
MTN3-like protein - Arabidopsis thaliana  
N/Alternate names: protein T8P19.250  
C/Species: Arabidopsis thaliana (mouse-ear creas)  
C/Date: 04-Feb-2000 #sequence\_revision 04-Feb-2000 #text\_change 09-Jul-2004

C:Accession: T46218  
R:Choiame, N.; Robert, C.; Brothier, P.; Wincker, P.; Catcollco, L.; Artiguenave, F.; Se submitted to the Protein Sequence Database, December 1999  
A:Reference number: Z23008  
A:Accession: T46218  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-289 <CHO>  
A:Cross-references: UNIPROT:Q9SMN5; EMBL:AL133315  
A:Experimental source: cultivar Columbia; BAC clone T8P19  
C:Genetics:  
A:Map position: 3  
A:introns: 18/1; 30/2; 100/3; 154/3; 194/3  
A>Note: T8P19.250

```

Query Match          46.0%; Score 622; DB 2; Length 289;
Best Local Similarity 47.9%; Pred. No. 9.2e-48;
Matches 126; Conservative 46; Mismatches 63; Indels 28; Gaps 4;
```

Oy	10	SFIFGLGNIVSFMWFLAPVPTFYKIKRKSEGYOAIPIYVAALFSAGLLIYYAYLRKA	69
Db	11	AFVGGLGNLISFAVFLSPVPTFYRIWKKTTEGQSIPYVAALFSATLMLYATOKKD	70
Oy	70	YLIVSINGFGCAIELTYISLFLEYPARKSKIET-GIMLLELGALGMVMPITYLAA	GS8
Db	71	FLVTINAFGCFIEITYISMFLAAPKPARMLTKMLLMNFSGFCAILLCQLVK	GAT
Oy	129	RVMIVGWICAINAVFAAPLSIRKOVIKTSVEFMPEPTSLPLTLCAITMFFYGFR	KD
Db	131	RAKIIGGICGVSVCFAPAPLSIRTVIKTSSVEFMPSLSLTITSAVIMLVGLAK	LD
Oy	189	FYIAFPNLGLFPGIVQMILYFYTK-----DSKRIDDK-----SDP-	225
Db	191	IYVAFPNNLGGPALQMLILYVVYKYCKTSPHLGEKEVEAKLPESLMDKLGT	VSSPE
Oy	226	----VREATSKSEGEVIITIED	244
Db	251	PISVVRQANKTCGNDRAEIED	273

RESULT 3  
T51837  
MTN3 homolog [imported] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 18-Aug-2000 #sequence\_revision 18-Aug-2000 #text\_change 09-Jul-2004  
C:Accession: T51837  
R:Cheong, J.J.  
submitted to the EMBL Data Library September 1998  
A:Description: An Arabidopsis cDNA clone encoding a protein homologous to Medicago truncatula  
A:Reference number: Z25480  
A:Accession: T51837  
A:Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: mRNA  
A:Residues: 1-285 <CHE>  
A:Cross-references: UNIPROT:O82587; EMBL:AF095641; PIDN:AAC64192.1  
C:Genetics:  
A>Note: MTN3

```

Query Match          45.7%; Score 618; DB 2; Length 285;
Best Local Similarity 47.1%; Pred. No. 2e-47;
Matches 123; Conservative 54; Mismatches 62; Indels 22; Gaps 4;
```

Oy	10	SFIFGLGNIVSFMWFLAPVPTFYKIKRKSEGYOAIPIYVAALFSAGLLIYYAYLRKA	69
Db	11	AFVGGLGNLISFAVFLSPVPTFYRICKKTTBEGFSIPYVAALFSATLMLYATOKKD	70
Oy	70	YLIVSINGFGCAIELTYISLFLEYPARKSKIET-GIMLLELGALGMVMPITYLAA	GS8
Db	71	FLVTINAFGCFIEITYISMFLAAPKPARMLTKMLLMNFSGFCAILLCQLVK	GAT
Oy	129	RVMIVGWICAINAVFAAPLSIRKOVIKTSVEFMPEPTSLPLTLCAITMFFYGFR	KD
Db	131	RAKIIGGICGVSVCFAPAPLSIRTVIKTSSVEFMPSLSLTITSAVIMLVGLAK	LD

[illegible]

A:Map position: 1

Query Match 25.9%; Score 350; DB 2; Length 202;

Best Local Similarity 37.2%; Pred. No. 8.8e-24; Matches 77; Conservative 40; Mismatches 82; Indels 8; Gaps 3;

QY 51 VALFSAGLLLYYA--YLKNAVLIVSINGFGCAIELTYISLFLVAPRKSKIFTGWLML 108

DB 1 MTLNCLSLMWGLPFPVSKNTLSTINGCAVIEETVYVILFLFYAPKKEKIKIFGFSC 60

QY 109 ELGALGMMPITVLLAEGSHRWMIIGWICAINVAFAPLSTIMROVYIKTSVEFMPFTL 168

DB 61 VLAVFAVVALVSLFALQGNRKLFCGLAATVFSIIIMVYASPSIMRLVTKSVFMPFPL 120

QY 169 SLFLTCATMWFPPYFGKDFYIAFPNIIIGFLGIYGMMLFYVYKDSKRIDDEKSDPVE 228

DB 121 SLFVPLCGISWPFYGLGRPFVAIPGFCALGTQLILFYCGNK--GEKS--AD 174

QY 229 ATKSEGVETIINIEDNSDALQSM 255

DB 175 AQKDEKSEVMKDEKKNVNGKDDLO 201

## RESULT 6

T04280

hypothetical protein F25124.60 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 30-Apr-1999 #sequence\_revision 30-Apr-1999 #text\_change 09-Jul-2004

C:Accession: T04280

R:Bevan, M.; Hilbert, H.; Braun, M.; Holzer, E.; Brandt, A.; Duesterhoeft, A.; Bancroft,

submitted to the Protein Sequence Database, March 1999

A:Reference number: 215261

A:Accession: T04280

A:Molecule type: DNA

A:Residues: 1-238 &lt;BEV&gt;

A:Cross-references: UNIPROT:Q9SN64; EMBL:AL049525

A:Experimental source: cultivar Columbia; BAC clone F25124

C:Genetics:

A:Map position: 4

A:Introns: 136/3; 176/3

A&gt;Note: F25124.60

Query Match 24.0%; Score 325; DB 2; Length 238;

Best Local Similarity 36.5%; Pred. No. 1.7e-21; Matches 74; Conservative 43; Mismatches 78; Indels 8; Gaps 5;

QY 30 PTFYKIYKRKSSBGYQAIIPYVALFSAAGLLLYYA--LRKNAVLIVSINGFGCAIELTYI 87

DB 7 PTFVRIYKKSVEYSPILYATLINCIVWLVGLPTVHPDSTLVITINGGILIEIVFL 66

QY 88 SLFLFAAPR-KSKIFTGWLMLLELGAIGMMPITVLLAEGSH-RWMIIGWICAINVAF 145

DB 67 TIFPVYCGRQKRLIISAVIAAEFAIILAVLVLTQHTTEKRTMSVGIKCCVFNNMY 126

QY 146 AAPLSIM--ROYIKTSVEFMPFTLSLFLTCATMWFPPYFGKDFYIAFPNIIIGFLG 202

DB 127 ASPLSVWVRKKNVITKSVFMPFWSVAGFLNAGVMTIYALMPDFPMALPBGICLFG 186

QY 203 IVQMLLY-FVYKDSKRIDDEKSD 224

DB 187 LAQLILGAYYKSTKRIMAREN 209

## RESULT 7

T01891

hypothetical protein F8M12.20 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 26-Feb-1999 #sequence\_revision 26-Feb-1999 #text\_change 09-Jul-2004

C:Accession: T01891

R:Madson, C.; Graves, T.; Cotton, M.; Modde, T.

submitted to the EMBL Data Library, July 1998

A:Description: The sequence of A. thaliana F8M12.

A:Reference number: 214450

A:Accession: T01891

A&gt;Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-249 &lt;MAD&gt;

A:Cross-references: UNIPROT:Q91628; EMBL:AF080118; NID:93513725; PID:93513744

A:Experimental source: cultivar Columbia

C:Genetics:

A:Map position: 4

A:Introns: 135/1; 147/3; 187/3

A&gt;Note: F8M12.20

Query Match 23.6%; Score 319.5; DB 2; Length 249;

Best Local Similarity 34.6%; Pred. No. 5.6e-21; Matches 74; Conservative 43; Mismatches 78; Indels 19; Gaps 5;

QY 30 PTFYKIYKRKSSBGYQAIIPYVALFSAAGLLLYYA--LRKNAVLIVSINGFGCAIELTYI 87

DB 7 PTFVRIYKKSVEYSPILYATLINCIVWLVGLPTVHPDSTLVITINGGILIEIVFL 66

QY 88 SLFLFAAPR-KSKIFTGWLMLLELGAIGMMPITVLLAEGSH-RWMIIGWICAINVAF 145

DB 67 TIFPVYCGRQKRLIISAVIAAEFAIILAVLVLTQHTTEKRTMSVGIKCCVFNNMY 126

QY 146 AAPLSIM-----ROYIKTSVEFMPFTLSLFLTCATMWFPPYFGKDFYI 191

DB 127 ASPLSVWVRKKNVITKSVFMPFWSVAGFLNAGVMTIYALMPDFPM 186

QY 192 APFNIIIGFLGIYGMMLY-FVYKDSKRIDDEKSD 224

DB 187 APNGICLFGILAQILLYGAYYKSTKRIMAREN 220

## RESULT 8

A71425

hypothetical protein - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

A:Variety: Columbia

C:Date: 03-Aug-1998 #sequence\_revision 03-Aug-1998 #text\_change 09-Jul-2004

C:Accession: A71425

R:Bevan, M.; Bancroft, I.; Bent, E.; Love, K.; Goodman, H.; Dean, C.; Bergkamp, R.; Dirki,

P.; Wedler, H.; Wedler, E.; Wamboldt, R.; Weltenssger, T.; Pohl, T.M.; Terry, N.; Gietz,

avanagh, T.; Hempel, S.; Kotler, P.; Entlan, K.D.; Rieger, M.; Schaeffer, M.; Funk, B.

Nature 391, 485-488, 1998

A:Authors: Mueller-Auer, S.; Silvey, M.; James, R.; Montfort, A.; Pons, A.; Puigdomenech,

erhoft, A.; Moores, T.; Jones, J.D.G.; Eneva, T.; Palme, K.; Benes, V.; Rechman, S.; Ansc,

C.; Chaiwattis, N.

A:Title: Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of Arabidopsis thal

A:Reference number: A71400; MUID:98121113; PMID:9461215

A:Accession: A71425

A&gt;Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-298 &lt;BEV&gt;

A:Cross-references: UNIPROT:Q23441; GB:297339; NID:92244901; PID:e327488; PID:92244949

C:Genetics:

A:Map position: 4COP9-4G3845

Query Match 21.5%; Score 291.5; DB 2; Length 298;

Best Local Similarity 32.7%; Pred. No. 2.1e-18; Matches 64; Conservative 36; Mismatches 49; Indels 47; Gaps 5;

QY 8 DLSFIFGLGNIVSPWFLAPV-----PTFYKIYKRKSSBGYQAIIPYV 51

DB 3 BASFYIGVIGNVSVLVSFLILRSILIIYOTKIRRETFMKIVGRSREYKSLPYIC 62

QY 52 ALFSAAGLLLYYALRKNAVLIVSINGFGCAIELTYISLFLVAPRKSKIFTGWLMLLELG 111

DB 63 TLIGSSILMTYTGIVTGEYLVSTVNGFALIVETIYVSLFLFYAPRLKLT-----VDVD 117

QY 112 A-LGMMPITVLLAEGS-----HRWMIIGWICAINVAFAPLSTIMROVITKTSVEFM 164

DB 118 AMNVVFPPIAAIYATSAFEDKMRQSIGFISAGNIIYSGPSLAM----- 165

QY 165 PFTLSLFLTCATMWF 180

Db 166 -----LTYCTTMY 174

## RESULT 9

T19016  
hypothetical protein C06G8.1 - Caenorhabditis elegans

C/Species: Caenorhabditis elegans

C/Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004

C/Accession: T19016

R/Matthews, P.  
submitted to the EMBL Data Library, March 1996

A/Reference number: Z19060

A/Accession: T19016

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-355 <full>

A/Cross-references: UNIPROT:Q17757; EMBL:Z70306; PIDD:CAA94322.1; GSPDB:GN00022; CESP:CC

A/Experimental source: clone C06G8

C/Genetics:

A/Gene: CESP:C06G8.1

A/Map position: 4

A/Introns: 31/2; 236/1

Query Match 18.0%; Score 243; DB 2; Length 355;

Best Local Similarity 27.9%; Pred. No. 5.1e-14;

Matches 72; Conservative 52; Mismatches 101; Indels 36; Gaps 12;

13 FGLIGNVSEFVFLAPVPTFY-----KIYKRSSSGYQAIPTMVALLFSAGLLLYAYL 65

9 FSLI-NLSTIAFTTGTGTFGCGIPICRQIMKRDITKISGAPLMGVGGCCMTTGYWL 67

66 RKNAYLVISINGSCAIELTYISLFLFY-APRSKIFGTGMLMLLGLG---MVPIT 120

68 -KNDGTVKWT--GCQY-IYTTYTIFYWCMTKKLV---ISLVLVGIIGCTSLVLAH 120

121 YLAEGRHWIVMICAIVAVFAAPLSIMROYITKSVEMPFITSLFLTCATMWP 180

121 FFGMKIRHP--IGIVCLITNIDFAAPLGIRVIRWATSTLPLCIANPLVSTWTF 177

181 FYGPFKDFYAFPIIIGFLFGIVQMLLYFYKDSKRIDDEKSDPV-----EATKSEK 235

178 LYGILKNDPFIIFNGVSLAPIQILFLV---PKPGORAPIVRLMIRGVVEET 234

236 VEIITNI-----EDDNDNALQ 252

235 KEIVAELEGCEDEKDKKGNRAQ 256

RESULT 10

A71607

Mtns/RAG1IP-like protein PFB0760w - malaria parasite (Plasmodium falciparum)

C/Species: Plasmodium falciparum

C/Date: 13-Nov-1998 #sequence\_revision 13-Nov-1998 #text\_change 09-Jul-2004

C/Accession: A71607

R/Gardner, M.O.; Tetteh, H.; Carucci, D.J.; Cummings, L.M.; Aravind, L.; Koonin, E.V.;

Perce, M.; Salzberg, S.; Zhou, L.; Sutton, G.G.; Clayton, R.; White, O.; Smith, H.O.

Science 282, 1126-1132, 1998

A/Title: Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum.

A/Reference number: A71600; MUID:99021743; PMID:9804551

A/Accession: A71607

A/Status: preliminary; nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1-686 <GAR>

A/Cross-references: UNIPROT:O96245; GB:AB001416; GB:AB001362; NID:G3845268; PIDD:AACT194

A/Experimental source: clone 3D7

C/Genetics:

A/Gene: PFB0760w

Query Match 17.3%; Score 233.5; DB 2; Length 686;

Best Local Similarity 26.5%; Pred. No. 7.3e-13;

Matches 65; Conservative 52; Mismatches 95; Indels 33; Gaps 8;

QY 23 MFLAPVPTFYKYIKYKRSSSGYQAIPTMVALLFSAGLLLYAYARKNAYLVISINGFCAT 82

Db 397 LIFD---PIIFKIIKKKTGELDPFYIILISFLVLVYGMILNNS-AIVFPVLGILIT 452

QY 83 ELTYISLFLFYAR---KSKIFT-----GMLMLELGAOMNPITLLAEGSRWIVG 134

Db 453 GILYCVIYHKCKNMMLKQKLSHYKICGFTCL-----LVAFLYILSYOYEV-FVG 504

QY 135 WICAAIVAVFAAPLSIMROYITKSVEMPFITSLFLTCATMWPFGFKDFYIAPF 194

Db 505 FVAFISIVNFGAPLSIYQIVIKKNSSLIPMEVTMGSLLCSFLMLTYGFTLDGFIIP 564

QY 195 NIIGFLFGIVQMLLYFY--KDSKRIDDEKSDPYREATKSEGVETI-----IN 241

Db 565 NLGCFILSLQVLIILILYSNKENTFNHSDITVSEISTKRNKXYIPDTNSMNFNEYN 624

QY 242 IEDDN 246

Db 625 VDEBN 629

RESULT 11

T26725

hypothetical protein Y39A1A.8 - Caenorhabditis elegans

C/Species: Caenorhabditis elegans

C/Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004

C/Accession: T26725

R/Mall, M.  
submitted to the EMBL Data Library, September 1998

A/Reference number: Z20257

A/Accession: T26725

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-250 <full>

A/Cross-references: UNIPROT:Q9XX26; EMBL:AL031633; PIDD:CAA21014.1; GSPDB:GN00021; CESP:Y

A/Experimental source: clone Y39A1A

C/Genetics:

A/Gene: CESP:Y39A1A.8

A/Map position: 3

A/Introns: 23/1; 52/1; 96/3; 152/3; 193/2

Query Match 16.3%; Score 220; DB 2; Length 250;

Best Local Similarity 31.7%; Pred. No. 3.8e-12;

Matches 58; Conservative 35; Mismatches 78; Indels 12; Gaps 6;

13 FGLIGNVSEFVFLAPVPTFYKYIKYKRSSSGYQAIPTMVALLFSAGLLLYAYARKNAYLI 72

Db 15 FTALSTVAF--FLCGIQTIRIKTRGSSBGTSPAPPLSLFSGGLFIQYGL-KDDVI 71

QY 73 VSINGFCATIELTYISLFLFYAPRKSKIFTGMLMLLGLAGMVPITVLLAE--GSH-- 128

Db 72 TYCNGIGCPLOACY--LMFYVYTRNRRLFNKYISIELGIIVV--YVVASTNSHLT 126

QY 129 RWMIVMICAIVAVFAAPLSIMROYITKSVEMPFITSLFLTCATMWPFGFKD 188

Db 127 KTTYGVGYCFLNIGSVAAFLFDIGKVRNKSSESLPLVCACFVVCLOMFEYIVDD 186

QY 189 FYI 191

Db 187 IVI 189

RESULT 12

JC4761

recombination activating gene 1 inducing protein - mouse

C/Species: Mus musculus (house mouse)

C/Date: 10-May-1996 #sequence\_revision 16-Aug-1996 #text\_change 09-Jul-2004

C/Accession: JC4761

R/Tagoh, H.; Kishi, H.; Muraguchi, A.

Biochem. Biophys. Res. Commun. 221, 744-749, 1996

A/Title: Molecular cloning and characterization of a novel stromal cell-derived cDNA enc

oding for.

A/Reference number: JC4761; MUID:96205347; PMID:8630032

A/Accession: J04761  
 A/Molecule type: mRNA  
 A/Residues: 1-221 <TAG>  
 C/Species: Caenorhabditis elegans  
 C/Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 09-Jul-2004  
 C/Comment: This protein is a membrane receptor involved in the induction of recombination  
 C/Genetics:  
 A/Gene: C2.3  
 C/Keywords: phosphoprotein  
 F/157/Binding site: phosphate (Thr) (covalent) #statue predicted

Query Match 16.8%; Score 216; DB 2; Length 221;  
 Best Local Similarity 26.0%; Pred. No. 7.6e-12;  
 Matches 47; Conservative 49; Mismatches 73; Indels 12; Gaps 3;

QY 37 KRKSSSEYQAIIPYVVALFSAAGLLLYAYARKNAVLYVINGFCATIEITYISLFLFYAPR 96  
 DB 35 RTSSVNDIQFLPPLTVDNNLSWLSVGLKGDGLII-VNSVGAIVLQTLVILYLAHSPQ 93  
 QY 97 KSKIFPGWMLLELALGALGMVPIITY-----LAEGRHWIVGVICAINVAFAAPLSI 151  
 DB 94 KHGV-----LQGTATLNAVLLGFGFVLLVDPLEARLQGLFCSVFPIISMVLSPLAD 147  
 QY 152 MROVITKTSVEFMPTLSLFLTCATWMPFYGFKKDYIAFPNIIIGFLFGIVQMLIYFV 211  
 DB 148 LAKIVQTKSTQRLSFSLITATLFCASWSIYGFRLRDPYIAVNPILPGLITSLIRLGLFCK 207  
 QY 212 Y 212  
 DB 208 Y 208

## RESULT 13

T31865  
 hypothetical protein C54F6.4 - Caenorhabditis elegans  
 C/Species: Caenorhabditis elegans  
 C/Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 09-Jul-2004  
 C/Accession: T31865  
 R/Giesel, C.; Bradshaw, H.  
 submitted to the EMBL Data Library, July 1997  
 A/Description: The sequence of C. elegans cosmid C54F6.  
 A/Reference number: Z21094  
 A/Accession: T31865  
 A/Molecule type: DNA  
 A/Residues: 1-233 <GCT>  
 A/Cross-references: UNIPROT:O16448; EMBL:AF016447; PIRN:AA065939.1; GSPDB:GN00023; CESP:  
 A/Experimental source: strain Bristol N2; clone C54F6  
 C/Genetics:  
 A/Gene: C54F6.4  
 A/Map position: 5  
 A/Intons: 15/2; 41/2; 82/3; 131/3; 171/3

Query Match 15.8%; Score 213.5; DB 2; Length 233;  
 Best Local Similarity 25.1%; Pred. No. 1.3e-11;  
 Matches 46; Conservative 46; Mismatches 84; Indels 7; Gaps 3;

QY 30 PTFYKTYKRSSEGYQAIIPYVVALFSAAGLLLYAYARKNAVLYVINGFCATIEITYISL 89  
 DB 16 PICLQIRQGVHDISGFPLMGTLVLPFWLRYGFLR-NVWMLISNCAIPIAV-FNAN 73  
 QY 90 FLFYAPRKSKIFPGWMLLELALGALGMVPIITYLLAEGSHRHWIVGVICAINVAFAAPL 149  
 DB 74 FLFYAPRKSKIFPGWMLLELALGALGMVPIITYLLAEGSHRHWIVGVICAINVAFAAPL 128  
 QY 150 SIKRQVITKTSVEFMPTLSLFLTCATWMPFYGFKKDYIAFPNIIIGFLFGIVQMLIY 209  
 DB 129 AGRVVLRDREVITLPEVLCLVQLVQCLMNLVIGILIQDFLAVPRAVGMISIVQLSLF 188  
 QY 210 FVY 212  
 DB 189 LIF 191

## RESULT 14

T32982  
 hypothetical protein K02D7.5 - Caenorhabditis elegans  
 C/Species: Caenorhabditis elegans  
 C/Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 09-Jul-2004  
 C/Accession: T32982  
 R/Du, Z.; Magg1, L.  
 submitted to the EMBL Data Library, February 1998  
 A/Description: The sequence of C. elegans cosmid K02D7.  
 A/Reference number: Z21259  
 A/Accession: T32982  
 A/Status: preliminary; translated from GB/EMBL/DBJ  
 A/Molecule type: DNA  
 A/Residues: 1-299 <DUZ>  
 A/Cross-references: UNIPROT:O45102; EMBL:AF045645; PIRN:AA02609.1; GSPDB:GN00023; CESP:  
 A/Experimental source: strain Bristol N2; clone K02D7  
 C/Genetics:  
 A/Gene: CESP:K02D7.5  
 A/Map position: 4  
 A/Intons: 25/2; 51/2; 92/3; 141/3; 181/3; 232/1

Query Match 15.5%; Score 209.5; DB 2; Length 299;  
 Best Local Similarity 27.3%; Pred. No. 4e-11;  
 Matches 57; Conservative 37; Mismatches 96; Indels 19; Gaps 4;

QY 24 VFAPVPTFYKTYKRSSEGYQAIIPYVVALFSAAGLLLYAYARKNAVLYVINGFCATIE 83  
 DB 20 LFFCGIPICMQIRQGVHDISGFPLMGVLSFPLRGGLKMD-YWIIYVNVGVACM 78  
 QY 84 LTYISLFLFYAPRKSKIFPGWMLLELALGALGMVPIITYLLAEGSHRHWIVGVICAINVA 143  
 DB 79 AFYCVFPLIYSLPK-KTFTCOLIVTSTIGWMLIAL-----KPNLDYLGVICMFENM 132  
 QY 144 VFAAPLSIKRQVITKTSVEFMPTLSLFLTCATWMPFYGFKKDYIAFPNIIIGFLFGI 203  
 DB 133 NFGAPLAGGVVLKQNEVSTLPMPKCVANFLVSSQWCLGNLVSDIYIIIPNGIGMFLAI 192  
 QY 204 VQMLTFYVYKDSKRIDEKSDPVEATKS 232  
 DB 193 VQMLTFVVL-----PIRENEKS 209

## RESULT 15

T32655  
 hypothetical protein K11D12.5 - Caenorhabditis elegans  
 C/Species: Caenorhabditis elegans  
 C/Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 09-Jul-2004  
 C/Accession: T32655  
 R/Henhaus, J.; Woldmann, P.; Gilling, B.  
 submitted to the EMBL Data Library, December 1997  
 A/Description: The sequence of C. elegans cosmid K11D12.  
 A/Reference number: Z21207  
 A/Accession: T32655  
 A/Status: preliminary; translated from GB/EMBL/DBJ  
 A/Molecule type: DNA  
 A/Residues: 1-224 <HEN>  
 A/Cross-references: UNIPROT:O44620; EMBL:AF039047; PIRN:AA094225.1; GSPDB:GN00023; CESP:  
 A/Experimental source: strain Bristol N2; clone K11D12  
 C/Genetics:  
 A/Gene: CESP:K11D12.5  
 A/Map position: 5  
 A/Intons: 48/2; 89/3; 141/3; 181/3

Query Match 10.8%; Score 146; DB 2; Length 224;  
 Best Local Similarity 20.5%; Pred. No. 1.3e-05;  
 Matches 45; Conservative 47; Mismatches 99; Indels 28; Gaps 8;

QY 22 FMVFLAPVPTFYKTYKRSSEGYQAIIPYVVALFSAAGLLLYAYARKNAVLYVINGFCG 79  
 DB 13 FSIQFTPLFMVFLDMHKGITADGSSVNFVLPMLVQSFMLRGRVNTNDQTNII-INSIN 71  
 QY 80 CAIEITYISLFLFYAPRKSKIFPGWMLLELALGALGMVPIITYLLAEGSHRHW-IVGWICA 138

Db 72 LVFPAYVSAFAYYQP-KRXYLIGQIV---AALAVKVAFAVVDTHDSASINDAMGSMMA 127  
QY 139 AINVAVFAAPLISIMRQVYIKTSVEFMP-FTLSLFTLCATWVFYGFKKDFYIAFPNI 196  
Db 128 GAQIFSLVSGIYEIKRAISMGTEYIPAGQFAIFTLI--LQWLLPGILHGNQPIAISNA 185  
QY 197 LGFLPGIVQMLLYFY-----KDSKRID 219  
Db 186 AGLVNIATLALYFFYPPLTWVPINFNIPQNKDAKVE 224

Search completed: December 24, 2004, 20:44:10  
Job time : 53 secs





```
Db      129 KRSLVIMICLVNI SVPAAPLCIMKRVIKTKSVSEFMPFSLTITNAVMWFFYGLLK 188
Qy      188 DFTYAFPNILGFLFGIYOMLLFYFYKDSKRIDDEKSPVEBATSKEGVEIINIIEDDN 246
Db      189 DFTYALPNTLGFILGFIQOMVLYLYRNA-----KPGLEBPTVQGLNHIIIDVAPN 241
```

## RESULT 2

```
US-10-425-115-214274
; Sequence 214274, Application US/10425115
; Publication No. US200400214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 214274
; LENGTH: 246
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_127016C.1.pep
US-10-425-115-214274
```

```
Query Match      49.5%; Score 669.5; DB 17; Length 246;
Best Local Similarity 53.8%; Pred. No. 8,7e-56;
Matches 120; Conservative 52; Mismatches 44; Indels 7; Gaps 2;

Qy      9  LSFIFGLGNIVSFMVFLAPVPTFYKTKYKRSRSGYQAIIPYVVALFSAGLLLYAYLRKN 68
Db      10  MAEPFGLGNIVSMTFLAPVPTFYKTKYKRSSTEGFSVPVVALFSAMLMIFYALKSN 69
Qy      69  AYLVISNGFCALIELTYISLFLFYAPRKSXIPTG-WMLLEGALGMVMPITYLAEGS 127
Db      70  ELLITINSAGCVETIYIMVLYAPRKAKLFTAKILLLNAGVFGVILLTLTLSSAQ 129
Qy      128  HRVITGMICAINVAAPAPLSTMRQVITKTSVEFMPFTLSFLTCATWMPFYGFKK 187
Db      130  RRVVLLGMVCVAFSVFVAPLSTIRQVTRTSVEFMPFSLSTLTVSAVVMFLYGLIK 189
Qy      188  DFTYAFPNILGFLFGIYOMLLFYFYKDS-----KRIDDEKSD 224
Db      190  DKYALPNTLGFISFGVQOMVLYLYRNAVTPRPVDFKVDADGKD 232
```

## RESULT 3

```
US-10-425-114-68239
; Sequence 68239, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 68239
; LENGTH: 299
; TYPE: PRT
; ORGANISM: Zea mays
```

```
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMFI0017025G08_FLI.pep
US-10-425-114-68239
Query Match      48.3%; Score 654; DB 15; Length 299;
Best Local Similarity 52.8%; Pred. No. 3.4e-54;
Matches 121; Conservative 46; Mismatches 50; Indels 12; Gaps 2;
```

```
Qy      10  SFIIFGLGNIVSFMVFLAPVPTFYKTKYKRSRSGYQAIIPYVVALFSAGLLLYAYLRKN 69
Db      7  AFAFGLGNIVSMTFLAPVPTFYKTKYKRSSTEGFSVPVVALFSAMLMIFYALKSN 66
Qy      70  YLVISNGFCALIELTYISLFLFYAPRKSXIPTGWL-LLELGALGMVMPITYLAEGSH 128
Db      67  TPLITINAGCVETIYIMVLYAPRKAKLFTAKIMALLNGVFGVILLTLTLFRGSK 126
Qy      129  RVMITGMICAINVAAPAPLSTMRQVITKTSVEFMPFTLSFLTCATWMPFYGFKKD 188
Db      127  RYVLLGMICVGFVSVPVAPLSTMRVITQTSVEFMPFSLSTLSAVVMFLYGLIKD 186
Qy      189  FTYAFPNILGFLFGIYOMLLFYFY-----KDSKRIDDEKSDPV 226
Db      187  KYVALPNTLGFIRGVQOMVLYLYRNKTPVAAAVGADAGLPSAADHV 235
```

## RESULT 4

```
US-10-425-115-214284
; Sequence 214284, Application US/10425115
; Publication No. US200400214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 214284
; LENGTH: 302
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_127025C.1.pep
US-10-425-115-214284
```

```
Query Match      48.3%; Score 653.5; DB 17; Length 302;
Best Local Similarity 50.4%; Pred. No. 3.8e-54;
Matches 122; Conservative 49; Mismatches 54; Indels 17; Gaps 3;
```

```
Qy      10  SFIIFGLGNIVSFMVFLAPVPTFYKTKYKRSRSGYQAIIPYVVALFSAGLLLYAYLRKN 69
Db      11  AFAFGLGNIVSMTFLAPVPTFYKTKYKRSSTEGFSVPVVALFSAMLMIFYALKSN 70
Qy      70  YLVISNGFCALIELTYISLFLFYAPRKSXIPTGWL-LLELGALGMVMPITYLAEGSH 128
Db      71  TPLITINAGCVETIYIMVLYAPRKAKLFTAKIMVLLNGVFGVILLTLTLFRGSK 130
Qy      129  RVMITGMICAINVAAPAPLSTMRQVITKTSVEFMPFTLSFLTCATWMPFYGFKKD 188
Db      131  RYVLLGMICVGFVSVPVAPLSTMRVITQTSVEFMPFSLSTLSAVVMFLYGLIKD 190
Qy      189  FTYAFPNILGFLFGIYOMLLFYFYKDSKRIDDEKSPVEBATSKEG-----VEIIT 240
Db      191  KYVALPNTLGFIRGVQOMVLYLYRN-----KTPVAATAGKADAGLPSAADHV 242
Qy      241  NI 242
Db      243  NI 244
```



RESULT 8  
US-10-425-115-214276  
Sequence 214276, Application US/10425115  
Publication No. US20040214272A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa, Thomas J.  
APPLICANT: Kovalic, David K.  
APPLICANT: Zhou, Yihua  
APPLICANT: Cao, Yongwei  
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
TITLE OF INVENTION: Plants  
FILE REFERENCE: 38-21(53222)B  
CURRENT APPLICATION NUMBER: US/10/425,115  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 369326  
SEQ ID NO 214276  
LENGTH: 302  
TYPE: PRT  
ORGANISM: Zea mays  
FEATURE:  
OTHER INFORMATION: Clone ID: MRT4577\_127018C.1.pcp  
US-10-425-115-214276

Query Match 48.1%; Score 650.5; DB 17; Length 302;  
Best Local Similarity 50.0%; Pred. No. 7.4e-54;  
Matches 121; Conservative 49; Mismatches 55; Indels 17; Gaps 3;  
QY 10 SFIFGLIGNIVSFMVFLAPVPTFYKTKRSSEGYOAIPIYVVALFSAAGLLIYAYLRKNA 69  
DB 11 AFAAGLIGNIVSFMVFLAPVPTFYKTKRSSEGYOAIPIYVVALFSAAGLLIYAYLRKNA 70  
QY 70 YLIIVSINGFGCAILETYISLFLFYAPRKSIFGTGMLM-LELGALGWMPTIYLLAEGSH 128  
DB 71 TELTINAGGVITITIVVMYFVAPKAKLFTAKIMALLNGVFGVILLTLTLFPGSK 130  
QY 129 RVMIVGICAINVAVAPLISIMROYIKTKSVEMFPTLSLFLTCATWMPFYGFPKD 188  
DB 131 RVVLGICVAVLSVAPLISIMROYIKTKSVEMFPTLSLFLTCATWMPFYGFPKD 190  
QY 189 FYIAFPNIGLFGIVQMLLYFYVKDSRIDEKSDPREATSKSGVEII 240  
DB 191 KYVALPVLGIFGVQMLLYFYVKDSRIDEKSDPREATSKSGVEII 242  
QY 241 NI 242  
DB 243 NI 244

RESULT 9  
US-10-424-599-142919  
Sequence 142919, Application US/10424599  
Publication No. US20040031072A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa, Thomas J.  
APPLICANT: Kovalic, David K.  
APPLICANT: Zhou, Yihua  
APPLICANT: Cao, Yongwei  
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with  
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
FILE REFERENCE: 38-21(53223)B  
CURRENT APPLICATION NUMBER: US/10/424,599  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 285684  
SEQ ID NO 142919  
LENGTH: 254  
TYPE: PRT  
ORGANISM: Glycine max  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT3847\_10006C.1.pcp  
US-10-424-599-142919

Query Match 48.0%; Score 649.5; DB 15; Length 254;  
Best Local Similarity 48.2%; Pred. No. 7.5e-54;  
Matches 121; Conservative 65; Mismatches 60; Indels 5; Gaps 4;  
QY 4 LRADDLSFIFGLIGNIVSFMVFLAPVPTFYKTKRSSEGYOAIPIYVVALFSAAGLLIYAYLRKNA 63  
DB 3 MHRBSMAFVFGVGNISLFGVFLAPVPTFYKTKRSSEGYOAIPIYVVALFSAAGLLIYAYLRKNA 62  
QY 64 YL-RKNAYLIVSINGFGCAILETYISLFLFYAPRKSIFGTGMLM-LELGALGWMPTIYLLAEGSH 121  
DB 63 FVARETALLIITINTFGIVVESIYLSFLFYARPKRPLTLTIKLLLNNGVFGAMLLSTL 122  
QY 122 LLAEGSHRVMIVGICAINVAVAPLISIMROYIKTKSVEMFPTLSLFLTCATWMPFYGFPKD 181  
DB 123 YLSKGAERLAIIGWICLVFNISVFAAPLFIIRVIVIKTSVEMFPTLSMFLTINAVMMPF 182  
QY 182 YGPFKDFYIAFPNIGLFGIVQMLLYFYVKDSRIDEKSDPREATSKSGVEII-1 240  
DB 183 YGLLRDYVALPVLGIFGVQMLLYFYVKDSRIDEKSDPREATSKSGVEII-1 240  
QY 241 NIEDNSDNL 251  
DB 241 GTMBEPNNGAV 251

RESULT 10  
US-10-437-963-170840  
Sequence 170840, Application US/10437963  
Publication No. US20040123343A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa, Thomas J.  
APPLICANT: Kovalic, David K.  
APPLICANT: Zhou, Yihua  
APPLICANT: Cao, Yongwei  
APPLICANT: Wu, Wei  
APPLICANT: Boukharov, Andrey A.  
APPLICANT: Barbazuk, Brad  
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with  
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
FILE REFERENCE: 38-21(53221)B  
CURRENT APPLICATION NUMBER: US/10/437,963  
CURRENT FILING DATE: 2003-05-14  
NUMBER OF SEQ ID NOS: 204966  
SEQ ID NO 170840  
LENGTH: 319  
TYPE: PRT  
ORGANISM: Oryza sativa  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (1) (319)  
OTHER INFORMATION: unsure at all Xaa locations  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT4530\_69127C.1.pcp  
US-10-437-963-170840

Query Match 47.9%; Score 648; DB 16; Length 319;  
Best Local Similarity 50.4%; Pred. No. 1.4e-53;  
Matches 116; Conservative 53; Mismatches 55; Indels 6; Gaps 2;  
QY 10 SFIFGLIGNIVSFMVFLAPVPTFYKTKRSSEGYOAIPIYVVALFSAAGLLIYAYLRKNA 69  
DB 12 AFAAGLIGNIVSFMVFLAPVPTFYKTKRSSEGYOAIPIYVVALFSAAGLLIYAYLRKNA 71  
QY 70 YLIIVSINGFGCAILETYISLFLFYAPRKSIFGTGMLM-LELGALGWMPTIYLLAEGSH 128  
DB 72 ELVVTINGVGVIEIYVLAIVAPKAKLFTAKIMALLNGVFGVILLTLTLFPGSK 131  
QY 129 RVMIVGICAINVAVAPLISIMROYIKTKSVEMFPTLSLFLTCATWMPFYGFPKD 188  
DB 132 RVVLGICVAVLSVAPLISIMROYIKTKSVEMFPTLSLFLTCATWMPFYGFPKD 191  
QY 189 FYIAFPNIGLFGIVQMLLYFYVKDSRIDEKSDPREATSKSGVEII 238

Db

192 VFVALLPNVLGFVFGVAQMALVMAVRSKKPLVASSSSAVVAA-----GLEI 236

RESULT 11  
US-10-767-701-46186  
; Sequence 46186, Application US/10767701  
; Publication No. US20040172684A1

Query Match	47.9%	Score 647.5	DB 16	Length 304
Best Local Similarity	50.9%	Pred No. 1.4e-53		
Matches 119, Conservative	49	Mismatches 65	Indels 1	Gaps 1

[illegible]

```
Qy      189 FYAAFPNIGLFGIVOMLLYFVVKDSKRIDDEKSDPVREATSKSGVEIIINI 242  
        | : | | | | | | | | | | : : : : |  
Db      191 KYVALPNILGTFGVGMVLTVLYMNKTPIVAVEAGDAGGKLPSAADENHLVNI 244
```

RESULT 12  
US-10-437-963-117614  
; Sequence 117614, Application US/10437963  
; Publication No. US20040123343A1

Query Match	47.8%	Score 646.5	DB 16	Length 303
Similarity	49.8%	Pred. No. 1.8e-53		
Best local				
Matches 119	Conservative 54	Mismatches 55	Indels 11	Gaps 2

[illegible]

```

RESULT 13
US-10-739-930
; Sequence 10939, Application US/10739930
; Publication NO. US20040216190A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
; TITLE OF INVENTION: PLANTS AND USES THEREOF FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-21(53377)B
; CURRENT APPLICATION NUMBER: US/10/739,930
; CURRENT FILING DATE: 2003-12-18
; NUMBER OF SEQ ID NOS: 11088
; SEQ ID NO 10939
; LENGTH: 294
; TYPE: PRT
; ORGANISM: Triticum aestivum
; FEATURE:
; OTHER INFORMATION: Clone ID: TRIAE-23APR03-C866_1.p
US-10-739-930-10939

```

Query Match	47.6%	Score 644.5	DB 17	Length 294
Best Local Similarity	51.3%	Pred. No. 2,7e-53		
Matches 116, Conservative	56	Mismatches 47	Indels 7	Gaps 2

```

QY 10 SFIFELNINIVSFWMLVLPVPFFPKYIKRKSSGEQALIPYMALFSAGLLIYAAVYARKNA 69
Db 11 AFAFELLGNIIVSEMYLAPLPFFYIKYISKSQIGQGSVYVVALPSAMLIYALLKSD 70
QY 70 YLIIVSINGFGCAIELTYISLFLFVAPRKSQIFTG-WLMLELGALGMVPIYLLAEGSH 128
Db 71 CLIIITNSAGCVIEIIYIIIVLTAPKQAKLFTAKILLILLNAGVFCLIIITLLTLLSEGEK 130
QY 129 RVMIVGMICAIINAVVAPAPLSIMIQVIKTSVEMFPFLTSLFLTICATMMFPYFGFKD 188
Db 131 RVMVMGMVCVGSVSVAFLAPLSVILVVRITRSEVEMFELSLSLTYSAAVWFLDYLKID 190
QY 189 FYIAEPNIIIGFLFGIVOMLIVFYKDS-----KRIIDKSDPVRE 228
Db 191 KYVALPNIILGFAPGVIOGLYALVNLNAPMSPAKQVDDADAIKVP 236

```

```

; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 117614
; LENGTH: 303
; TYPE: prt
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_21002C.1.pcp
; US-10-437-963-117614

RESULT 14
US-10-425-115-214279
; Sequence 214279, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

```





```
RESULT 2
; US-08-902-853-3
; Sequence 3, Application US/08902853
; Patent No. 5945330
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Corley, Neil C.
; APPLICANT: Shah, Puri
; APPLICANT: Lal, Preeti
; TITLE OF INVENTION: HUMAN LONGEVITY-ASSURANCE PROTEIN HOMOLOGS
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/902,853
; FILING DATE: Herewith
; CLASSIFICATION: ?
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0245 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 380 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: HMT2NOT01
; CLONE: 493014
; US-08-902-853-3

Query Match 7.6%; Score 107.5; DB 2; Length 380;
Best Local Similarity 27.6%; Pred. No. 0.0044;
Matches 70; Conservative 28; Mismatches 109; Indels 47; Gaps 14;

QY 1 MAQLRADLSPFIRGLGNIVSFWFLAPVPTFYIKRKSEGV---QATP-----YMW 52
DB 130 LKFRERASMRTEFLA-FLAGAAVVDKPFYDM--KKWEGPPIOSTLPSQYWMIE 186
QY 53 L-ESAGLLLYAYAKKAAVIVSINGCAIELTYISLFLPYAPRKSKIFGWMLELG 111
DB 187 LSPFWSLLFSLASDVKKDKPEQIHHVATITLISFWMFANY-----IRAG 232
QY 112 ALGWMV--PTTYLAESSHRYMIVGICAAINV-AVPAAPLSIRQVIKTKSVEMFPTL 168
DB 233 TLIMALHDSDDYLL-ESAKMFNYAGMKNTCNNIFIVAFITRLVI-----LPFWI 284
QY 169 SLFTTLGATWVFGEFKDQFYIAFPNIGL-----FGIVOMLLYVYKDSKRIDEKS 223
DB 285 -LHCTLVYPELVPAFGYFFPNSMGLVQLLHFMAYLILRMHKPI--TGKLVEDERS 341
QY 224 DPVEATKSKGV 237
DB 342 D--RETTESSEGE 353
```

```
RESULT 3
US-09-270-767-40416
; Sequence 40416, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 40416
; LENGTH: 140
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-40416

Query Match 7.6%; Score 103.5; DB 4; Length 140;
Best Local Similarity 27.7%; Pred. No. 0.003;
Matches 26; Conservative 19; Mismatches 40; Indels 9; Gaps 1;

QY 133 VGMICAINVAAPAPLSIRQVIKTKSVEMFPTLSLFTLCAIWMFFYGFKKDQFYIA 192
DB 45 LGMLITGILVMWVGSPDLHLPKTIERKSTEGMPFPIIFAGNLVAFSWTLTAISIKTWV 104
QY 193 FPNILGFLFGIVOMLLYFVY-----KDSKR 217
DB 105 LQNLILLVLGGIQLSMFAIYPNKPAAEKPKDSK 138

RESULT 4
US-09-270-767-55632
; Sequence 55632, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 55632
; LENGTH: 140
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-55632

Query Match 7.6%; Score 103.5; DB 4; Length 140;
Best Local Similarity 27.7%; Pred. No. 0.003;
Matches 26; Conservative 19; Mismatches 40; Indels 9; Gaps 1;

QY 133 VGMICAINVAAPAPLSIRQVIKTKSVEMFPTLSLFTLCAIWMFFYGFKKDQFYIA 192
DB 45 LGMLITGILVMWVGSPDLHLPKTIERKSTEGMPFPIIFAGNLVAFSWTLTAISIKTWV 104
QY 193 FPNILGFLFGIVOMLLYFVY-----KDSKR 217
DB 105 LQNLILLVLGGIQLSMFAIYPNKPAAEKPKDSK 138

RESULT 5
US-09-248-796A-26637
; Sequence 26637, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstein et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
```



;;  
CURRENT FILING DATE: 1999-02-12  
PRIOR APPLICATION NUMBER: US 60/074,725  
PRIOR FILING DATE: 1998-02-13  
PRIOR APPLICATION NUMBER: US 60/096,409  
PRIOR FILING DATE: 1998-08-13  
NUMBER OF SEQ ID NOS: 28208  
SEQ ID NO 26937  
LENGTH: 252  
TYPE: PRT  
ORGANISM: Candida albicans  
US-09-248-796A-26937

Query Match 7.5%; Score 101; DB 4; Length 252;  
Best Local Similarity 24.0%; Pred. No. 0.012;

Matches 56; Conservative 37; Mismatches 98; Indels 42; Gaps 10;

QY 43 GYQAIIPYVAL-ESAGL--LYYAYLRKNAYLIVSINGFGCAIELTYISLFLYAPRKSX 99  
DB 17 GAQSVFCIVVLGASGLADGVNDRVTPALVSI-----LNLIFYSYILLMPTILK 70  
QY 100 IFPGMLLELGLGMMPTTYLLAGSHRVMVGMCAINAVFAAPLSIMROYIKK 159  
DB 71 NFSPSVIIL--VAEFIFVFYSAMGALAAVIPSQSGC--DYGSSYSAKSLIKALI--- 122  
QY 160 SVEFMPTLSLFTLCATMWFYGF-----FKDFYIAPNLIPLFGIVQMLLY 209  
DB 123 -----PFTLFNMLFATSEFLGYSFIPQVSSRGFSITLPA-----REFGAIPTDPA 172  
QY 210 FVYKDSKRIDEKSD-PVREATKSKGVE---IINIEDNSDNLQSMED 257  
DB 173 LPFGKKYAVDPVTDAIANAEVTGSGIENDAPKVASVGDNEATVGLASSEED 225

RESULT 6  
US-09-199-737-2  
Sequence 2, Application US/09199737A  
Patent No. 6287788  
GENERAL INFORMATION:  
APPLICANT: Bard, Jonathan A.  
APPLICANT: Borowsky, Beth  
APPLICANT: Smith, Kelli E.  
APPLICANT: Branchek, Theresa A.  
APPLICANT: Gerald, Christophe P.G.  
APPLICANT: Jones, Kenneth A.  
TITLE OF INVENTION: DNA Encoding Galanin GALR3 Receptors And Uses Thereof  
FILE REFERENCE: 52241-D-PCT-US  
CURRENT APPLICATION NUMBER: US/09/199,737A  
CURRENT FILING DATE: 1998-11-25  
NUMBER OF SEQ ID NOS: 59  
SOFTWARE: Patentin Ver. 2.0 - beta  
SEQ ID NO 2  
LENGTH: 370  
TYPE: PRT  
ORGANISM: Rat  
US-09-199-737-2

Query Match 7.5%; Score 101; DB 3; Length 370;  
Best Local Similarity 21.6%; Pred. No. 0.021;

Matches 46; Conservative 40; Mismatches 67; Indels 60; Gaps 10;

QY 9 LSFIFGLGNIVSFMVFLAPVPTFYKIYKRKSESG-----YQAIIPYVAL 53  
DB 26 LIFLLGMVNGVLAVLLOPGPS---AWQEPSSITDPLFINLAVADLCFLCCVPOQAI 82  
QY 54 FSAGLLLYAYLRKNAYLIVSINGFGCAIELTYISLFLYA-----PRSKIF 101  
DB 83 YTLDAWLFQAFVCKTCHLITLITMYASSFTLAANVSLDRYLARHPLRSRALRTPRARAA 142  
QY 102 TG--WMLLELGLGMMPTTY--LLAEGSHRVMIVGMICA--AINAVFAA-----PLS 150  
DB 143 VGLVWL---LALFSAPLYSTYGTAVYVYGALELCVPMEDARRALDVATFAAGYLLPVA 198  
QY 151 IMROYIKTSVEFMPTLSLFTLCATMWFYGF 183

DB 199 VV-----SLAYGRTLGC-FLWAAVG 216

RESULT 7  
US-08-900-230-2  
Sequence 2, Application US/08900230  
Patent No. 6329197  
GENERAL INFORMATION:

APPLICANT: Bard, Jonathan A.  
TITLE OF INVENTION: DNA ENCODING GALANIN GALR3 RECEPTORS AND  
TITLE OF INVENTION: USES THEREOF

NUMBER OF SEQUENCES: 59  
CORRESPONDENCE ADDRESS:

ADDRESSER: Cooper & Dunham LLP  
STREET: 1185 Avenue of the Americas

CITY: New York  
STATE: New York

COUNTRY: U.S.A.  
ZIP: 11036

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/900,230

FILING DATE: 23-JUL-1997

CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:

NAME: White, John P.  
REGISTRATION NUMBER: 28,678

REFERENCE/DOCKET INFORMATION: 52241-C/JPW/ADM  
TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-278-0400  
TELEFAX: 212-391-0525

INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:

LENGTH: 370 amino acids  
TYPE: amino acid

STRANDEDNESS: single  
MOLECULE TYPE:

TOPOLOGY: linear  
HYPOTHEICAL: NO

ANTI-SENSE: NO  
US-08-900-230-2

Query Match 7.5%; Score 101; DB 3; Length 370;  
Best Local Similarity 21.6%; Pred. No. 0.021;

Matches 46; Conservative 40; Mismatches 67; Indels 60; Gaps 10;

QY 9 LSFIFGLGNIVSFMVFLAPVPTFYKIYKRKSESG-----YQAIIPYVAL 53  
DB 26 LIFLLGMVNGVLAVLLOPGPS---AWQEPSSITDPLFINLAVADLCFLCCVPOQAI 82  
QY 54 FSAGLLLYAYLRKNAYLIVSINGFGCAIELTYISLFLYA-----PRSKIF 101  
DB 83 YTLDAWLFQAFVCKTCHLITLITMYASSFTLAANVSLDRYLARHPLRSRALRTPRARAA 142  
QY 102 TG--WMLLELGLGMMPTTY--LLAEGSHRVMIVGMICA--AINAVFAA-----PLS 150  
DB 143 VGLVWL---LALFSAPLYSTYGTAVYVYGALELCVPMEDARRALDVATFAAGYLLPVA 198  
QY 151 IMROYIKTSVEFMPTLSLFTLCATMWFYGF 183  
DB 199 VV-----SLAYGRTLGC-FLWAAVG 216

RESULT 8  
US-09-058-333A-2  
Sequence 2, Application US/09058333A  
Patent No. 6368812  
GENERAL INFORMATION:

```

; APPLICANT: Bard, Jonathan A
; APPLICANT: Borowsky, Beth
; APPLICANT: Smith, Kelli E
; TITLE OF INVENTION: DNA ENCODING GALANIN GALR3 RECEPTORS
; TITLE OF INVENTION: AND USES THEREOF
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION NUMBER:
; APPLICATION NUMBER: US/09/058,333A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 52241-E/JPW/KDB
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 278 0400
; TELEFAX: 212 391 0525
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 370 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-058-333A-2

Query Match
Best Local Similarity 7.5%; Score 101; DB 3; Length 370;
Matches 46; Conservative 40; Mismatches 67; Indels 60; Gaps 10;

QY 9 LSPFGLGNIVSFMVFLAPVTFYKIKYKRSSEG-----YQALPYWAL 53
Db 26 LIFLGNVGNGLVLAVALQPEPS--AMQEPSTTDLPILNLAVALDCLICVFPQAI 82
QY 54 FSAGLIYYAVLRKNAVLIVSINGFGCAIETLYISLFFVA-----PRKSKIF 101
Db 83 YTLDAWLFQAFVCKTVHLLIYLTMYASSFTLAASLDRIYLAVRHPLRSRALRTNRRAA 142
QY 102 TG--WMLLELGAIGWMPITY--LLAEGSHRVMIVGICA--AINAVFPA-----PLS 150
Db 143 VGLVWL-----LAAIFSAPIYLSYGVTVRYGALCLCPAMEDARRRDLVATFPAAGYLLPVA 198
QY 151 IMROVITKSVEMFPTLSLFTLCATMWFYFG 183
Db 199 VV-----SLAYGRITLC-FLMAAVG 216

RESULT 9
US-09-328-352-5497
; Sequence 5497, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; NUMBER OF SEQ ID NOS: 1999-06-04
; SEQ ID NO 5497
; LENGTH: 462
; TYPE: PRT
```

```

; ORGANISM: Acinetobacter baumannii
; US-09-328-352-5497

Query Match
Best Local Similarity 7.4%; Score 100.5; DB 4; Length 462;
Matches 56; Conservative 42; Mismatches 92; Indels 57; Gaps 13;

QY 11 FIFGLGNIVSFMVFLAPVTFYKIKYKRSSEGYQALPYWALP-----SA 56
Db 224 FFIGALLAIYVFRIRGLLET--QSPFKNAQAEFDQPKSGFALFKYPKAEFTVLETLG 281
QY 57 GLLIYYAVLRKNAVLIVSINGFGC--AIELTYISLFFVAPRKSIFTGWL-----M 106
Db 282 GTLAFTYTTTYLQKYLWNTSGFTKPBPAQTITTLALPFMCLOP--LAGALSRIGRKPL 338
QY 107 LLELGAIGWMPITYLLAE--GSHRVMIVGIC-----AAINAVFAPAPLSIMR 153
Db 339 MIAFGYTGVL--FTYILFDLTANTHMYTRAFWLCLGGLIMVTGYTSINAVKA---ELFP 393
QY 154 QVITKSVEMFPTLSLFTLCATMWFYGFYK--KQFYIAFPNIGLFGIYQIMLY 209
Db 394 AHIRALGVA-LPYAIAINTL-FGGLAEFFALSPKAEAGHESWFIYVIMIF-----ISLLY 447
QY 210 FVYKDSK 216
Db 448 IFMKDKT 454

RESULT 10
US-09-543-681A-5282
; Sequence 5282, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709,1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 5282
; LENGTH: 287
; TYPE: PRT
; ORGANISM: Proteus mirabilis
; US-09-543-681A-5282

Query Match
Best Local Similarity 7.3%; Score 99; DB 4; Length 287;
Matches 47; Conservative 35; Mismatches 85; Indels 78; Gaps 7;

QY 18 NIVSFMVFLAPVTFYKIKYKRSSEGYQALPYW---VALFSAGLLIYYAVLRKNAVLIVS 74
Db 28 SIYAVCAVLSLVFTIIPSLIRKQNGYRGIALIPAVVALYTHAISLKFIFRPHSGONLS 87
QY 75 INFGCAIETLYISLFFVAPRKSIFGTWMLLELGAIGWMPITYLLAEGSHRVMIVG 134
Db 88 LTNLGAIVSLMVCVINTIYASR-----GRAMFLPIYCYCS-----IVN 126
QY 135 WTCATN-----VAVFAAPLSIMROVITKSVEM 164
Db 127 LIIAALPGHFVTHLESSTSLFIHIGIALISYATLLIALLYALQSLWLDYQLNKKTKES 186
QY 165 P-----FTLS-----LFTLCATMWFYGFYKDYIAFPNIGLFGYVQW 206
Db 187 PQMPPLMSIERKMFHTYGVVLLTLTLCTGLGLMDNIFGE-----NIHKSIFSIIAW 240
QY 207 LLYFV 211
Db 241 FVYII 245
```

RESULT 11  
US-09-134-001C-3972  
Sequence 3972, Application US/09134001C  
Patent No. 6380370  
GENERAL INFORMATION:  
APPLICANT: Lynn Doucette-Stamm et al  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS  
FILE REFERENCE: GTC-007  
CURRENT APPLICATION NUMBER: US/09/134,001C  
CURRENT FILING DATE: 1998-08-13  
PRIOR APPLICATION NUMBER: US 60/064,964  
PRIOR FILING DATE: 1997-11-08  
PRIOR APPLICATION NUMBER: US 60/055,779  
PRIOR FILING DATE: 1997-08-14  
NUMBER OF SEQ ID NOS: 5674  
SEQ ID NO 3972  
LENGTH: 381  
TYPE: PRT  
ORGANISM: Staphylococcus epidermidis  
US-09-134-001C-3972

Query Match 7.3%; Score 99; DB 3; Length 381;  
Best Local Similarity 20.3%; Pred. No. 0.035;  
Matches 52; Conservative 40; Mismatches 80; Indels 84; Gaps 12;  
QY 15 LGNITSPM-----VFLAPVPTFYKIKRKSEGYQAIPLYWALFSA-----56  
DB 168 ILSNVSYSITSTFISTVFLWVFPIIYMKDHEK--IPAIKFKGKERKVFVDLKD 225  
QY 57 -----GLLYVAVLRKNAVLIIVSINGGCAIELTYISLFLFYAPRK 97  
DB 226 LNTLSYIOGQVTSIIIGIILYIGY-----TITGPTPLVLVAGVA 270  
QY 98 SKI-FTG-WMLLELALGAV-MPTVYLLAEGSHRVWICAMINAVFAAPL--SIW 152  
DB 271 NLIPFLGPWLSFPAAILGIIDGPSTFI-----WVCV--VTLIAQQLGNAV 315  
QY 153 ROVIKRSVEFMPTLSPLTLCATMWFYGFPRKQPIYAFNPLIGFLFIQVMLLYFY 212  
DB 316 TPNVMKSSISIHPLTIIIVLAAIDGGLF-----TILVAVP-----LVAVIKTLVSNIF 365  
QY 213 KDSKRIIDKSDPVRE 228  
DB 366 KYRQRIYDKNANSVKD 381

RESULT 12  
US-09-134-000C-5441  
Sequence 5441, Application US/09134000C  
Patent No. 6617156  
GENERAL INFORMATION:  
APPLICANT: Lynn Doucette-Stamm et al  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
FILE REFERENCE: 032796-032  
CURRENT APPLICATION NUMBER: US/09/134,000C  
CURRENT FILING DATE: 1998-08-13  
PRIOR APPLICATION NUMBER: US 60/055,778  
PRIOR FILING DATE: 1997-08-15  
NUMBER OF SEQ ID NOS: 6812  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 5441  
LENGTH: 395  
TYPE: PRT  
ORGANISM: Enterococcus faecalis  
US-09-134-000C-5441

Query Match 7.3%; Score 99; DB 4; Length 395;  
Best Local Similarity 23.7%; Pred. No. 0.037;  
Matches 49; Conservative 35; Mismatches 93; Indels 30; Gaps 8;  
QY 18 NIVFNVFLAPVPTFYKIKRKSEGYQAIPLYM-----VALFSAGLLLYVAVLRKNAVLIY 73

DB 3 SVVKTLLGLGFIPSW-----RKKKEHDWLPMMFGWTFMFRS-YVLYEIIIRGTAFFLI 56  
QY 74 SINGRGCAIELTYISLF-LFYAPRKSKITFTGMLLELALGAMWPIIT-----YLLAG 126  
DB 57 PLAPFG-----LFALFYEKERRLNGLTFPIFLVAVLTYYGITAIRTNVFLITL 107  
QY 127 SRWMTVGICAMINAVFAAPLSIMROYIKTSVEFMPTLSPLTLCATMWFYGFPRK 186  
DB 108 AIIAIIATITALLAKRYALITIFLYNNAVVVMKRSRLNLTLLATISLTLLIYNF- 166  
QY 187 KDFYIAFPNIGLFLGIQVMLP-YFVY 212  
DB 167 --FQSLPFWLSPLTIAPIFLTYFAF 191

RESULT 13  
US-09-711-164-409  
Sequence 409, Application US/09711164  
Patent No. 6589738  
GENERAL INFORMATION:  
APPLICANT: Forsyth, R. Allyn  
APPLICANT: Ohlsen, Kari  
APPLICANT: Zyskind, Judith  
TITLE OF INVENTION: GENES ESSENTIAL FOR MICROBIAL PROLIFERATION AND ANTISENSE THERET  
FILE REFERENCE: ELITRA.008A  
CURRENT APPLICATION NUMBER: US/09/711,164  
CURRENT FILING DATE: 2000-11-09  
PRIOR APPLICATION NUMBER: US 60/164415  
PRIOR FILING DATE: 1999-11-9  
NUMBER OF SEQ ID NOS: 469  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 409  
LENGTH: 567  
TYPE: PRT  
ORGANISM: Escherichia coli  
US-09-711-164-409

Query Match 7.3%; Score 98.5; DB 4; Length 567;  
Best Local Similarity 20.2%; Pred. No. 0.068;  
Matches 56; Conservative 48; Mismatches 86; Indels 87; Gaps 14;  
QY 23 MYFLAPVPTFYKIKRKSEGYQAIPLYWALF---SAGLLY---AYLRKNAVLIIVSIN 76  
DB 1 MAVSERELIVKRIKPSKIDFTLAPFYISIFPLCSGLFQYFPAFTKGCY-----54  
QY 77 GFCAIELTYISLFLFYAPRKSKITFTGMLLELALGAMWPIYLLAEGSHRV-----130  
DB 55 --ECSWKLDYIKQFYF-----SLETAWYL--ISAVAFIASVFI---QRIKAYLT 99  
QY 131 MYGWCALAI-----NAVFAAPLSIMROYIKTSVEFMPTLSPLTLCATMWF 180  
DB 100 LAITWVLTITDVALIHADNITAMNNIILNITLNL-----FGAILSLFNCLSNLI 152  
QY 181 FVGFPKQPIYAFNPLIGFLFIQVML-----LYVYKDSKRIIDKSDPVREART 231  
DB 153 HANKIKH-----IPILSMIPVLSAIIIAIITAVIYLF-----188  
QY 232 SKGVEIINIIEDNSDNA---LQSEKDFSLRSTK 265  
DB 189 AROAVEIEMDI-SEGSDIAYGVKDNESGFPLNDK 224

RESULT 14  
US-09-328-352-7553  
Sequence 7553, Application US/09328352  
Patent No. 6562958  
GENERAL INFORMATION:  
APPLICANT: Gary L. Breton et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
FILE REFERENCE: GTC99-03PA  
CURRENT APPLICATION NUMBER: US/09/328,352



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 24, 2004, 18:58:30 ; Search time 118 Seconds  
(without alignments)  
805.621 Million cell updates/sec

Title: US-09-743-885A-1  
Perfect score: 1353  
Sequence: 1 MAQLRADLISFIFGLGNIV.....NSDNLQSMEXDFSLRTSK 265

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 35872929 residues  
Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_23Sep04:\*

- 1: geneseqp1980a:\*
- 2: geneseqp1990a:\*
- 3: geneseqp2000a:\*
- 4: geneseqp2001a:\*
- 5: geneseqp2002a:\*
- 6: geneseqp2003a:\*
- 7: geneseqp2003b:\*
- 8: geneseqp2004a:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1353	100.0	265	3	AAVS8647
2	1353	100.0	265	3	AAV44803
3	704.5	52.1	258	3	AAAG43420
4	703.5	52.0	258	3	AAAG4273
5	623.5	46.1	292	3	AAAG06230
6	622	46.0	289	3	AAAG54308
7	622	46.0	289	3	AAAG6337
8	622	46.0	289	8	ADN73901
9	621.5	45.9	289	3	AAAG9770
10	621.5	45.9	289	3	AAAG1971
11	618	45.7	285	3	AAAG1683
12	615.5	45.5	289	7	ABMT3930
13	615	45.5	268	7	ABMT3783
14	590	43.6	294	3	AAAG42299
15	564.5	41.7	281	3	AAAG42959
16	564.5	41.7	304	3	AAAG07674
17	488.5	36.1	195	3	AAAG3421
18	488.5	36.1	195	3	AAAG4274
19	447	33.0	247	8	ADM47668
20	447	33.0	247	8	ADM47960
21	433.5	32.0	234	3	AAAG9771
22	433.5	32.0	234	3	AAAG1972
23	431.5	31.9	231	3	AAAG1973
24	431.5	31.9	231	3	AAAG9772
25	431	31.9	228	3	AAAG1684

26	429	31.7	291	8	ADM47961	Polypepti
27	427.5	31.6	235	3	AAAG06231	Arabidops
28	411	30.4	239	3	AAAG42300	Arabidops
29	403.5	29.8	251	3	AAAG51964	Arabidops
30	401.5	29.7	232	3	AAAG42301	Arabidops
31	400.5	29.6	251	3	AAAG19711	Arabidops
32	399.5	29.5	226	3	AAAG42960	Arabidops
33	399.5	29.5	249	3	AAAG07675	Arabidops
34	395	29.2	300	7	ABM74461	DNA clone
35	383.5	28.3	258	3	AAAG16725	Arabidops
36	363.5	26.9	239	3	AAAG46827	Arabidops
37	363.5	26.9	239	3	AAAG11624	Arabidops
38	353.5	26.1	225	3	AAAG22336	Arabidops
39	353.5	26.1	236	3	AAAG22335	Arabidops
40	350	25.9	202	3	AAAG17669	Arabidops
41	337	24.9	200	3	AAAG54309	Arabidops
42	337	24.9	200	3	AAAG26338	Arabidops
43	328.5	24.3	203	3	AAAG06232	Arabidops
44	323.5	23.9	185	3	AAAG41685	Arabidops
45	318.5	23.5	242	7	ABM73820	DNA clone

## ALIGNMENTS

RESULT 1  
AAVS8647  
ID AAVS8647 standard; protein; 265 AA.

AAVS8647;  
12-SEP-2003 (revised)  
11-APR-2000 (first entry)  
Petunia nectary-specific NECT1 protein.  
NECT1; nectary; nectar; transgenic plant; honey.  
Petunia x hybrida.  
EP974667-A1.  
26-JAN-2000.

16-JUL-1998; 98EP-00202375.  
16-JUL-1998; 98EP-00202375.  
16-JUL-1998; 98EP-00202375.

(CPRO-) CPRO-DLO CENT PLANTENVERDELINGS REPROD.

Creemers J, Angenent GC, Kater MM;  
WPI; 2000-108400/10.  
N-PSDB; AA235493.

Novel DNA sequences used to produce modified honey, the metabolites of which can be isolated and purified.

Claim 1; Page 15; 56pp; English.

The present sequence is that of Petunia hybrida NECT1 protein that is highly expressed in the nectaries of petunia and weakly expressed in the stamens. The sequence was deduced from cDNA (see AA235493) obtained by differential display. The present invention provides a method for producing recombinant proteins in honey. The honey is manufactured by insects, preferably honeybees, that collect the nectar of transgenic plants. The NECT1 gene and its promoter (see AA235496) can be utilised in expression cassettes for the production of transgenic plants that produce a protein of interest in their nectar. The function of NECT1 has not yet been determined. (Updated on 12-SEP-2003 to standardise OS field)

Sequence 265 AA;

5

Query Match 100.0%; Score 1353; DB 3; Length 265;  
Best Local Similarity 100.0%; Pred. No. 9,9e-151;  
Matches 265; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAQLRADDSIFIFGLGNIVSFVWFLAPVPTFYKIKRKSSSEGYQAIPIYVVALFSAGLL 60  
DB 1 MAQLRADDSIFIFGLGNIVSFVWFLAPVPTFYKIKRKSSSEGYQAIPIYVVALFSAGLL 60

QY 61 YYAYLKRNAYLIVSINGFGCAIELTYISLFLYAPRKSIFFTGWLMLBELGALGMVPI 120  
DB 61 YYAYLKRNAYLIVSINGFGCAIELTYISLFLYAPRKSIFFTGWLMLBELGALGMVPI 120

QY 121 YLLAEGSHRVMIVGICAAINVAFAAPLSIMRQVITKTSVEFMPFTLSFLTLCATMP 180  
DB 121 YLLAEGSHRVMIVGICAAINVAFAAPLSIMRQVITKTSVEFMPFTLSFLTLCATMP 180

QY 181 FYGPFKKDFIAPFNITLGFIVQMLLYFYVYKDSKRIDEKSDPYREATKSKGEVEIII 240  
DB 181 FYGPFKKDFIAPFNITLGFIVQMLLYFYVYKDSKRIDEKSDPYREATKSKGEVEIII 240

QY 241 NIEDDNSDNALQSMEXKDFSLRLRTSK 265  
DB 241 NIEDDNSDNALQSMEXKDFSLRLRTSK 265

## RESULT 2

AAV44803  
ID AAV44803 standard; protein; 265 AA.

AC AAV44803;

DT 12-SEP-2003 (revised)  
DT 17-MAY-2000 (first entry)

XX Petunia hybrida nectary-specific protein NECL.

KM NECL; nectary-specific protein; metabolite; recombinant protein;  
KM transgenic plant; nectar; honey; insect; bee; pharmaceutical; enzyme;  
KM biotest; antioxidant; food additive.

XX Petunia x hybrida.

PN W0200004176-A1.

PD 27-JAN-2000.

PF 15-JUL-1999; 99MO-NL000453.

PR 16-JUL-1998; 98EP-00202375.

PR 14-DEC-1998; 98EP-00204215.

PA (CPR0-) CPR0-DLO CENIT PLANTENVERDEDELINGS REPROD.

PI Creemers J, Angenent GC, Kater MM;

DR WPI, 2000-182438/16.

DR N-PSDB; AAZ50200.

PT New DNA encoding Petunia hybrida nectary-specific proteins, useful for,  
e.g. producing modified honey.

PS Claim 1; Page 39; 93pp; English.

XX The present sequence is a nectary-specific protein NECL which was  
CC isolated from nectaries of Petunia hybrida strain W15 using mRNA  
CC differential display system. NECL resembles membrane protein and is  
CC strongly expressed in the nectaries of P. hybrida. A DNA sequence from  
CC the promoter region upstream of nectary-specific expressed sequence e.g.  
CC NECL and FBp15 DNAs is used in a recombinant DNA construct comprising a  
CC DNA encoding a metabolite preferably recombinant protein, a DNA encoding  
CC a signal peptide that targets the recombinant protein to the nectar and  
CC optionally a signal sequence functional in plants for the transcription  
CC termination and polyadenylation of an RNA molecule. The DNA construct is

CC useful for producing transgenic plants which excrete recombinant proteins  
CC in its nectar. The nectar is processed into honey by insects (preferably  
CC bees) and the desired protein is easily recovered from it. The  
CC recombinant proteins are useful for pharmaceutical purposes, as enzymes  
CC for biotests and antioxidants for food additives. (Updated on 12-SEP-2003  
CC to standardise OS field)

XX Sequence 265 AA;

Query Match 100.0%; Score 1353; DB 3; Length 265;  
Best Local Similarity 100.0%; Pred. No. 9,9e-151;  
Matches 265; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAQLRADDSIFIFGLGNIVSFVWFLAPVPTFYKIKRKSSSEGYQAIPIYVVALFSAGLL 60  
DB 1 MAQLRADDSIFIFGLGNIVSFVWFLAPVPTFYKIKRKSSSEGYQAIPIYVVALFSAGLL 60

QY 61 YYAYLKRNAYLIVSINGFGCAIELTYISLFLYAPRKSIFFTGWLMLBELGALGMVPI 120  
DB 61 YYAYLKRNAYLIVSINGFGCAIELTYISLFLYAPRKSIFFTGWLMLBELGALGMVPI 120

QY 121 YLLAEGSHRVMIVGICAAINVAFAAPLSIMRQVITKTSVEFMPFTLSFLTLCATMP 180  
DB 121 YLLAEGSHRVMIVGICAAINVAFAAPLSIMRQVITKTSVEFMPFTLSFLTLCATMP 180

QY 181 FYGPFKKDFIAPFNITLGFIVQMLLYFYVYKDSKRIDEKSDPYREATKSKGEVEIII 240  
DB 181 FYGPFKKDFIAPFNITLGFIVQMLLYFYVYKDSKRIDEKSDPYREATKSKGEVEIII 240

QY 241 NIEDDNSDNALQSMEXKDFSLRLRTSK 265  
DB 241 NIEDDNSDNALQSMEXKDFSLRLRTSK 265

## RESULT 3

AAG43420  
ID AAG43420 standard; protein; 258 AA.

AC AAG43420;

DT 18-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 54268.

KM Protein identification; signal transduction pathway; metabolic pathway;  
KM hybridisation assay; genetic mapping; gene expression control; promoter;  
KM termination sequence.

XX Arabidopsis thaliana.

PN BP1033405-A2.

PD 06-SEP-2000.

PF 25-FEB-2000; 2000EP-00301439.

PR 25-FEB-1999; 99US-0121825P.

PR 05-MAR-1999; 99US-0123180P.

PR 09-MAR-1999; 99US-0123548P.

PR 23-MAR-1999; 99US-0125788P.

PR 25-MAR-1999; 99US-0126264P.

PR 29-MAR-1999; 99US-0126785P.

PR 01-APR-1999; 99US-0127462P.

PR 06-APR-1999; 99US-0128234P.

PR 08-APR-1999; 99US-0128714P.

PR 16-APR-1999; 99US-0129845P.

PR 19-APR-1999; 99US-0130077P.

PR 21-APR-1999; 99US-0130449P.

PR 23-APR-1999; 99US-0130510P.

PR 23-APR-1999; 99US-0130891P.

PR 28-APR-1999; 99US-0131449P.

PR 30-APR-1999; 99US-0132048P.

PR 30-APR-1999; 99US-0132407P.

PR 04-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 06-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132487P.  
PR 11-MAY-1999; 99US-0132486P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144684P.  
PR 21-JUL-1999; 99US-0144681P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 26-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147415P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0158293P.  
PR 13-OCT-1999; 99US-0158294P.  
PR 13-OCT-1999; 99US-0158294P.  
PR 13-OCT-1999; 99US-0158295P.  
PR 14-OCT-1999; 99US-0158329P.  
PR 14-OCT-1999; 99US-0158330P.  
PR 14-OCT-1999; 99US-0158331P.  
PR 14-OCT-1999; 99US-0158637P.  
PR 14-OCT-1999; 99US-0158638P.  
PR 18-OCT-1999; 99US-0158584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.

Query	Match	Score	DB	Length
Query	Best Local Similarity	52.1%;	Score 704.5; DB 3;	length 258;
Matches	136; Conservative	48; Mismatches	62; Indels	1; Gaps
Db	1	2	3	4
Qy	4	5	6	7
Db	3	4	5	6
Qy	64	65	66	67
Db	63	64	65	66
Qy	123	124	125	126
Db	123	124	125	126
Qy	183	184	185	186
Db	183	184	185	186
Qy	243	244	245	246
Db	243	244	245	246
RESULT 4				
AAAG24273				
ID	AAAG24273	standard; protein; 258 AA.		
AC	AAAG24273;			
AC	AAAG24273;			
DT	17-OCT-2000	(first entry)		
DE	Arabidopsis thaliana protein fragment SEQ ID NO: 27884.			
XX	Arabidopsis thaliana.			
XX	Protein identification; signal transduction pathway; metabolic pathway;			
XX	hybridisation assay; genetic mapping; gene expression control; promoter;			
XX	termination sequence.			
XX	Arabidopsis thaliana.			
XX	EP1033405-A2.			
XX	06-SEP-2000.			
XX	25-FEB-2000; 2000EP-00301439.			
XX	25-FEB-1999; 99US-0121825P.			
XX	05-MAR-1999; 99US-0123180P.			
XX	09-MAR-1999; 99US-0123548P.			
XX	23-MAR-1999; 99US-0125788P.			
XX	25-MAR-1999; 99US-0126264P.			
XX	29-MAR-1999; 99US-0126785P.			
XX	01-APR-1999; 99US-0127462P.			
XX	06-APR-1999; 99US-0128234P.			

PR	08-APR-1999;	99US-01268741P
PR	16-APR-1999;	99US-01268454P
PR	19-APR-1999;	99US-01300777P
PR	21-APR-1999;	99US-01305449P
PR	23-APR-1999;	99US-01305510P
PR	23-APR-1999;	99US-01308911P
PR	28-APR-1999;	99US-01314449P
PR	30-APR-1999;	99US-01320488P
PR	04-MAY-1999;	99US-01324407P
PR	04-MAY-1999;	99US-01324644P
PR	05-MAY-1999;	99US-01324659P
PR	06-MAY-1999;	99US-01324877P
PR	07-MAY-1999;	99US-01328631P
PR	11-MAY-1999;	99US-01342565P
PR	14-MAY-1999;	99US-01342181P
PR	14-MAY-1999;	99US-01342191P
PR	14-MAY-1999;	99US-01342219P
PR	18-MAY-1999;	99US-01343701P
PR	18-MAY-1999;	99US-01347688P
PR	19-MAY-1999;	99US-01349414P
PR	20-MAY-1999;	99US-01351142P
PR	21-MAY-1999;	99US-01353533P
PR	24-MAY-1999;	99US-01353639P
PR	25-MAY-1999;	99US-01360211P
PR	27-MAY-1999;	99US-01363732P
PR	28-MAY-1999;	99US-01367829P
PR	01-JUN-1999;	99US-01372222P
PR	03-JUN-1999;	99US-01375588P
PR	04-JUN-1999;	99US-01375502P
PR	07-JUN-1999;	99US-01377744P
PR	08-JUN-1999;	99US-01380949P
PR	10-JUN-1999;	99US-01385407P
PR	10-JUN-1999;	99US-01384847P
PR	14-JUN-1999;	99US-01391119P
PR	16-JUN-1999;	99US-01394522P
PR	16-JUN-1999;	99US-01394523P
PR	16-JUN-1999;	99US-01394533P
PR	17-JUN-1999;	99US-01394542P
PR	18-JUN-1999;	99US-01394945P
PR	18-JUN-1999;	99US-01394555P
PR	18-JUN-1999;	99US-01394565P
PR	18-JUN-1999;	99US-01394613P
PR	18-JUN-1999;	99US-01394623P
PR	18-JUN-1999;	99US-01394633P
PR	18-JUN-1999;	99US-01397508P
PR	18-JUN-1999;	99US-01397633P
PR	21-JUN-1999;	99US-01398171P
PR	22-JUN-1999;	99US-01398999P
PR	23-JUN-1999;	99US-01400533P
PR	23-JUN-1999;	99US-01400354P
PR	24-JUN-1999;	99US-01406939P
PR	24-JUN-1999;	99US-01408233P
PR	29-JUN-1999;	99US-01440991P
PR	30-JUN-1999;	99US-01412877P
PR	01-JUL-1999;	99US-01418424P
PR	01-JUL-1999;	99US-01421215P
PR	02-JUL-1999;	99US-01420555P
PR	06-JUL-1999;	99US-01423309P
PR	08-JUL-1999;	99US-01428033P
PR	09-JUL-1999;	99US-01429270P
PR	12-JUL-1999;	99US-01429777P
PR	13-JUL-1999;	99US-01435424P
PR	15-JUL-1999;	99US-01436244P
PR	15-JUL-1999;	99US-01440050P
PR	16-JUL-1999;	99US-01440868P
PR	16-JUL-1999;	99US-01440868P
PR	19-JUL-1999;	99US-01443255P
PR	19-JUL-1999;	99US-01443311P
PR	19-JUL-1999;	99US-01443322P



PR	13-OCT-1999;	99US-0144333P
PR	19-JUL-1999;	99US-0144334P
PR	19-JUL-1999;	99US-0144335P
PR	20-JUL-1999;	99US-0144332P
PR	20-JUL-1999;	99US-0144884P
PR	21-JUL-1999;	99US-0144814P
PR	21-JUL-1999;	99US-0145086P
PR	21-JUL-1999;	99US-0145088P
PR	22-JUL-1999;	99US-0145087P
PR	22-JUL-1999;	99US-0145089P
PR	22-JUL-1999;	99US-0145192P
PR	23-JUL-1999;	99US-0145145P
PR	23-JUL-1999;	99US-0145218P
PR	23-JUL-1999;	99US-0145224P
PR	26-JUL-1999;	99US-0145276P
PR	27-JUL-1999;	99US-0145913P
PR	27-JUL-1999;	99US-0145918P
PR	27-JUL-1999;	99US-0145919P
PR	28-JUL-1999;	99US-0145951P
PR	28-JUL-1999;	99US-0146386P
PR	02-AUG-1999;	99US-0146388P
PR	02-AUG-1999;	99US-0146389P
PR	03-AUG-1999;	99US-0147038P
PR	04-AUG-1999;	99US-0147204P
PR	05-AUG-1999;	99US-0147302P
PR	05-AUG-1999;	99US-0147192P
PR	05-AUG-1999;	99US-0147260P
PR	06-AUG-1999;	99US-0147303P
PR	06-AUG-1999;	99US-0147416P
PR	09-AUG-1999;	99US-0147453P
PR	09-AUG-1999;	99US-0147935P
PR	10-AUG-1999;	99US-0148171P
PR	11-AUG-1999;	99US-0148311P
PR	12-AUG-1999;	99US-0148341P
PR	13-AUG-1999;	99US-0148555P
PR	13-AUG-1999;	99US-0148684P
PR	16-AUG-1999;	99US-0149368P
PR	17-AUG-1999;	99US-0149415P
PR	18-AUG-1999;	99US-0149426P
PR	20-AUG-1999;	99US-0149723P
PR	20-AUG-1999;	99US-0149929P
PR	23-AUG-1999;	99US-0149902P
PR	23-AUG-1999;	99US-0149930P
PR	25-AUG-1999;	99US-0150566P
PR	26-AUG-1999;	99US-0150884P
PR	27-AUG-1999;	99US-0151065P
PR	27-AUG-1999;	99US-0151066P
PR	27-AUG-1999;	99US-0151080P
PR	30-AUG-1999;	99US-0151303P
PR	31-AUG-1999;	99US-0151318P
PR	01-SEP-1999;	99US-0151930P
PR	07-SEP-1999;	99US-0152363P
PR	10-SEP-1999;	99US-0153070P
PR	13-SEP-1999;	99US-0153758P
PR	15-SEP-1999;	99US-0154018P
PR	16-SEP-1999;	99US-0154039P
PR	20-SEP-1999;	99US-0154779P
PR	22-SEP-1999;	99US-0155113P
PR	23-SEP-1999;	99US-0155486P
PR	24-SEP-1999;	99US-0155659P
PR	28-SEP-1999;	99US-0156458P
PR	29-SEP-1999;	99US-0156569P
PR	04-OCT-1999;	99US-0157117P
PR	05-OCT-1999;	99US-0157753P
PR	06-OCT-1999;	99US-0157865P
PR	07-OCT-1999;	99US-0158029P
PR	08-OCT-1999;	99US-0158232P
PR	12-OCT-1999;	99US-0158359P
PR	13-OCT-1999;	99US-0159293P
PR	13-OCT-1999;	99US-0159294P

PR	13-OCT-1999;	99US-0159295P.
PR	14-OCT-1999;	99US-0159329P.
PR	14-OCT-1999;	99US-0159330P.
PR	14-OCT-1999;	99US-0159331P.
PR	14-OCT-1999;	99US-0159637P.
PR	14-OCT-1999;	99US-0159638P.
PR	18-OCT-1999;	99US-0159584P.
PR	21-OCT-1999;	99US-0160741P.
PR	21-OCT-1999;	99US-0160767P.
PR	21-OCT-1999;	99US-0160768P.
PR	21-OCT-1999;	99US-0160770P.
PR	21-OCT-1999;	99US-0160814P.
PR	21-OCT-1999;	99US-0160815P.
PR	22-OCT-1999;	99US-0160980P.
PR	22-OCT-1999;	99US-0160981P.
PR	22-OCT-1999;	99US-0160989P.
PR	25-OCT-1999;	99US-0161404P.
PR	25-OCT-1999;	99US-0161405P.
PR	25-OCT-1999;	99US-0161406P.
PR	26-OCT-1999;	99US-0161359P.
PR	26-OCT-1999;	99US-0161360P.
PR	26-OCT-1999;	99US-0161361P.
PR	28-OCT-1999;	99US-0161920P.
PR	28-OCT-1999;	99US-0161922P.
PR	28-OCT-1999;	99US-0161933P.
PR	29-OCT-1999;	99US-0162142P.

Query Match	52.0%	Score 703.5;	DB 3;	Length 258;
Best Local Similarity	55.1%	Pred. No. 4.9e-74;		
Matches 136;	Conservative 48;	Mismatches 62;	Indels 1;	Gaps 1;

Oy		4	LRRADLFFIFFLGLGNIVSFMVFLAPPTFYKITYGRKSSEGOAIPYMWALFSAGLLLYYA	63
Xx		:	:::::	:      :
Dd		3	LQVEINAFNLGCLGNISFGVFSLSPVTFTFCIGYKKSSKGQSPIYCALSATILLYYG	62
Oy		64	YLKNNAVLVISINGPGCAIEITYSLEFLFAPRKSKIFTGMML-ELGALGWMPITYL	122
Xx		:	:::::	:      :
Dd		63	IMKTHAVLIISINFSGFIEIISTYLETILYAPREAKSTLKLIYICNIGGIGLLILVNLT	122
Oy		123	LAEGSHVMIMGICALINVAVPAAPISIMROVKTKTSVEEFPRTLSDLTLTCATMPPFY	182
Xx		:	:::::	:      :
Dd		123	LVPQHNVSYVGWCCAYLSLVAFPASPLSVMRKVIKTKTSVEETMPTLSLTLNMVMMPFY	182
Oy		183	GFEFKDFVIAPENILGFPLGVQLLYPVYKDSCRIDEDKSDPYREATKSEGEVIIINI	242
Xx		:	:::::	:      :
Dd		183	GLLIKDXFIAMPNILGFLFGVAQMILYMMYOGSTKTDLPTENOANANKTDVNEVDIVAVEL	242
Oy		243	EDDNSDN	249
Xx				:
Dd		243	PDVGS DN	249
RESULT 5 AAG06230 ID AAG06230 standard; protein: 292 AA.				
Xx	AC	AAG06230;		
Xx	DT	17-OCT-2000	(first entry)	
Xx	DE	Arabidopsis thaliana protein fragment SEQ ID NO: 2928.		
KX	Protein identification: signal transduction pathway; metabolic pathway; hybridisation assay; genetic mapping; gene expression control; promoter; termination sequence.			
Xx	OS	Arabidopsis thaliana.		
Xx	PM	EPI033405-A2.		
Xx	PD	06-SEP-2000.		
PF	25-FEB-2000; 2000EP-00301439.			

XX  
PR 25-FEB-1999; 99US-0121825P.  
PR 05-MAR-1999; 99US-0123180P.  
PR 09-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 28-APR-1999; 99US-0130891P.  
PR 30-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 04-APR-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 07-MAY-1999; 99US-0132487P.  
PR 11-MAY-1999; 99US-0132863P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140655P.  
PR 28-JUN-1999; 99US-0140833P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 25-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145513P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 10-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149175P.  
PR 17-AUG-1999; 99US-0149426P.  
PR 18-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150684P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151382P.  
PR 01-SEP-1999; 99US-0151383P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.





XX AC AAG26337;  
XX DT 17-OCT-2000 (first entry)  
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 30756.  
XX KW Protein identification; signal transduction pathway; metabolic pathway;  
XX hybridization assay; genetic mapping; gene expression control; promoter;  
XX termination sequence.  
XX OS Arabidopsis thaliana.  
XX PN EP1033405-A2.  
XX PD 06-SEP-2000.  
XX PF 25-FEB-2000; 2000EP-00301439.  
XX PR 25-FEB-1999; 99US-0121825P.  
PR 05-MAR-1999; 99US-0123180P.  
PR 09-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 28-APR-1999; 99US-0130891P.  
PR 30-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 04-MAY-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 06-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135535P.  
PR 24-MAY-1999; 99US-0135529P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145226P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145919P.  
PR 02-AUG-1999; 99US-0145951P.  
PR 04-AUG-1999; 99US-0146386P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0148684P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.



Db 11 AEFVGLGLNLISFAVFLSPVPTFYRIWKKTTEGFSQISIPYVALFSATLMLYVAQCKDV 70  
QY 70 YIVVINGRCALIELTYISLFLFYAPRKSIFT-GWLMLELGAQWNPITYLLAESGH 128  
Db 71 FLVLTINAGCELETITYISMFLAYAPKPMMLTVKMLMNFQFCALLCOFLVKAT 130  
QY 129 RWIMVGMICAAINVAVFAAPLSIMROVITKSVFEMPTLSLFLTLCAWTFEYGFPKD 188  
Db 131 RAKITGICVGSVCFFAPLSIRIVITRSVEMWPFSLTLTISAVIMLYGLAKD 190  
QY 189 FTIAPNNITGLFGIVOMLLFYVK-----DSKRIDEK-----SDP- 225  
Db 191 IVVAFPMVGFALGMLQVIVVYKCKTSPHLGEKEVAALKPEVSLMLKLGTVSSPE 250  
QY 226 ----VREATSKESGEIINIED 244  
Db 251 PISVVRQANKCTCGNDRAEIED 273

RESULT 9  
AAG49770  
ID AAG49770 standard; protein; 289 AA.  
XX AAG49770;  
AC AAG49770;  
XX 18-OCT-2000 (first entry)  
DT 18-OCT-2000 (first entry)  
XX Arabidopsis thaliana protein fragment SEQ ID NO: 62997.  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 62997.  
XX Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridization assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX Arabidopsis thaliana.  
OS Arabidopsis thaliana.  
XX EP1033405-A2.  
PN EP1033405-A2.  
XX 06-SEP-2000.  
PD 06-SEP-2000.  
XX 25-FEB-2000; 2000EP-00301439.  
PF 25-FEB-2000; 2000EP-00301439.  
XX 25-FEB-1999; 99US-0121825P.  
PR 05-MAR-1999; 99US-0123180P.  
PR 09-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129645P.  
PR 19-APR-1999; 99US-0130077P.  
PR 23-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 28-APR-1999; 99US-0130891P.  
PR 30-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 30-APR-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 06-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135533P.  
PR 24-MAY-1999; 99US-0135629P.

PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138647P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 02-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 26-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.





PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.

PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145919P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149262P.  
PR 20-AUG-1999; 99US-0149712P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151338P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-015263P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157853P.  
PR 06-OCT-1999; 99US-0157855P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159310P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.

PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161922P.  
PR 28-OCT-1999; 99US-0161933P.  
PR 29-OCT-1999; 99US-0162142P.  
  
Query Match 45.9%; Score 621.5; DB 3; Length 289;  
Best Local Similarity 53.4%; Pred. No. 2.8e-64;  
Matches 110; Conservative 46; Mismatches 49; Indels 1; Gaps 1;  
  
QY 9 LSFIFGLGNIVSMFLAVPFPFPIYKTKKSGQAIYPMALFSAGLLYAYAVRK 68  
Db 8 LAFVFGILGNIIISFVCLAPIPTFVKIYKKSSEGVSIPIYSLFSAMLMYAMTKD 67  
  
QY 69 AYLVISNGFCAIELTYISLFLFYAPRKSKIPT-GWMLLELGCALGMWPIITLLAEGS 127  
Db 68 AMMLITNSFAFVQVQIYISLFFFYAPKKEKTLTVKVLVDVDFGFAIFVLTITIHAN 127  
  
QY 128 HRVMIWGITCAINNAVFAAPLPSIMROVITKSVEMPFLLSLFTLCATMFPFYPFKK 187  
Db 128 KRVOVGCIYIMVAPLSYFVAPLGIIRKVIKTKSAEFMPGLSFFLLTSAVMWFYGLLK 187  
  
QY 188 DFYIAPFNIGFPGIYQMLTFYK 213  
Db 188 DNNIAPNVLGPIFGVQMLTFYK 213  
  
RESULT 11  
AAG41683  
ID AAG41683 standard; protein; 285 AA.  
AC AAG41683;  
XX 18-OCT-2000 (first entry)  
XX Arabidopsis thaliana protein fragment SEQ ID NO: 51893.  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 51893.  
XX  
XX Protein identification; signal transduction pathway; metabolic pathway;  
KM hybridisation assay; genetic mapping; gene expression control; promoter;  
XX termination sequence.  
XX  
XX Arabidopsis thaliana.  
OS  
XX  
XX EPI033405-A2.  
PN  
XX  
XX 06-SEP-2000.  
PD  
XX  
PF 25-FEB-2000; 2000EP-00301439.  
XX  
XX 25-FEB-1999; 99US-0121825P.  
PR 05-MAR-1999; 99US-0123180P.  
PR 09-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129645P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 28-APR-1999; 99US-0130891P.  
PR 30-APR-1999; 99US-0132048P.  
PR 04-MAY-1999; 99US-0132407P.  
PR 05-MAY-1999; 99US-0132484P.  
PR 06-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 07-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132663P.

PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135333P.  
PR 21-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142820P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.

PR	23-JUL-1999	99US-01452218
PR	23-JUL-1999	99US-01452246
PR	26-JUL-1999	99US-01452766
PR	27-JUL-1999	99US-01459138
PR	27-JUL-1999	99US-01459199
PR	27-JUL-1999	99US-01459519
PR	28-JUL-1999	99US-01459519
PR	28-JUL-1999	99US-01459519
PR	02-AUG-1999	99US-01463866
PR	02-AUG-1999	99US-01463866
PR	02-AUG-1999	99US-01463866
PR	02-AUG-1999	99US-01463866
PR	02-AUG-1999	99US-01463866
PR	03-AUG-1999	99US-01470388
PR	03-AUG-1999	99US-01470388
PR	04-AUG-1999	99US-01472046
PR	04-AUG-1999	99US-01473026
PR	05-AUG-1999	99US-01771826
PR	05-AUG-1999	99US-01472606
PR	06-AUG-1999	99US-01473036
PR	06-AUG-1999	99US-01474166
PR	09-AUG-1999	99US-01474833
PR	09-AUG-1999	99US-01474833
PR	10-AUG-1999	99US-01479356
PR	10-AUG-1999	99US-01481716
PR	11-AUG-1999	99US-01483119
PR	12-AUG-1999	99US-01483416
PR	13-AUG-1999	99US-01485656
PR	13-AUG-1999	99US-01485656
PR	16-AUG-1999	99US-01493686
PR	17-AUG-1999	99US-01494175
PR	18-AUG-1999	99US-01494466
PR	20-AUG-1999	99US-01497226
PR	20-AUG-1999	99US-01497226
PR	20-AUG-1999	99US-01497236
PR	20-AUG-1999	99US-01499296
PR	23-AUG-1999	99US-01499026
PR	23-AUG-1999	99US-01499306
PR	26-AUG-1999	99US-01505666
PR	26-AUG-1999	99US-01505846
PR	27-AUG-1999	99US-01510656
PR	27-AUG-1999	99US-01510666
PR	27-AUG-1999	99US-01510666
PR	27-AUG-1999	99US-01510666
PR	30-AUG-1999	99US-01513036
PR	31-AUG-1999	99US-01514836
PR	01-SEP-1999	99US-01519336
PR	01-SEP-1999	99US-01519336
PR	07-SEP-1999	99US-01523636
PR	10-SEP-1999	99US-01530706
PR	13-SEP-1999	99US-01537586
PR	15-SEP-1999	99US-01540186
PR	16-SEP-1999	99US-01540396
PR	20-SEP-1999	99US-01547796
PR	22-SEP-1999	99US-01551336
PR	23-SEP-1999	99US-01554866
PR	24-SEP-1999	99US-01556596
PR	24-SEP-1999	99US-01565486
PR	28-SEP-1999	99US-01565966
PR	29-SEP-1999	99US-01571176
PR	04-OCT-1999	99US-01577536
PR	05-OCT-1999	99US-01577536
PR	06-OCT-1999	99US-01578656
PR	07-OCT-1999	99US-01580226
PR	08-OCT-1999	99US-01582326
PR	12-OCT-1999	99US-01583656
PR	13-OCT-1999	99US-01592836
PR	13-OCT-1999	99US-01592846
PR	13-OCT-1999	99US-01593296
PR	14-OCT-1999	99US-01593316
PR	14-OCT-1999	99US-01593316
PR	14-OCT-1999	99US-01596376
PR	14-OCT-1999	99US-01596386
PR	18-OCT-1999	99US-01595846
PR	21-OCT-1999	99US-01607416
PR	21-OCT-1999	99US-01607676
PR	21-OCT-1999	99US-01607766
PR	21-OCT-1999	99US-01607766
PR	21-OCT-1999	99US-01608146
PR	21-OCT-1999	99US-01608156
PR	22-OCT-1999	99US-01609806

PR	22-Oct-1999	99US-0160981P
PR	22-Oct-1999	99US-0160981P
PR	25-Oct-1999	99US-0161404P
PR	25-Oct-1999	99US-0161405P
PR	25-Oct-1999	99US-0161406P
PR	26-Oct-1999	99US-0161359P
PR	26-Oct-1999	99US-0161360P
PR	26-Oct-1999	99US-0161361P
PR	26-Oct-1999	99US-0161920P
PR	28-Oct-1999	99US-0161982P
PR	28-Oct-1999	99US-0161982P
PR	29-Oct-1999	99US-0162142P
PR	29-Oct-1999	99US-0162143P

Query Match	45.7%	Score 618	DB 3	Length 285
Best Local Similarly	47.1%	Pred. No. 7e-64		
Matches 123	Conservative 54	Mismatches 62	Indels 22	Gaps 4

```
Qy 10 SFIFGLGNIVSFMVFLAPVPTFYKIKRKSSEGVAIPYMWALFSAGLLLYAYLRQNA 69
      :::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 11 AFVFGLLGNLISFAVFLSPVPTFYRCCKKTTGFSQIPIYVVALFSMLMLVYATQKKDV 70
```

```
Qy      70 YLIVSINGPGCAIELTYISLFPLVAPRKSKFT-GWIMLLLELCALGMVPITYLLAESGSH 1298  
       :|::||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
Db      71 FLATINSFGCFIETIIYISIFVAFASKKARMLTVKLLILMRGGFCILLILOFLAKGT 1306
```

```

QY      129 RMTVGWICAINVAFAAPISIMROVIKTKYSVEENPFTLSLTLCATMMFEYGFPEKD 188
      | : | | : | | | | | : | | | | | | | | : | | |
Db      131 RAKITGICGCFGSVCYFAAPLSTIRVVIKTKYSVEYNPFSLSLTITSAVIMLLYGLAKD 190

```

```

Qy 189 FYIAFPNIGFLPGIVOMLIFYVK-----DSKRIDDEKSDPVREAT-KSKE 234
      |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 191 IYVAFPNVIGFVLGALQMLIVVYKCKTSPDLVEKELEAAKLPEVSI DMVKGLTSPB 250

```

```

QY      235 GVEIIT-----NIEDDNSD 248
          | | : | | | :
Db      251 PVAITVRSVNTCNCNDRAE 271

```

RESULT 12  
ABM73930  
ID ABM73930 standard; protein; 289 AA

AC	ABM73930;
XX	
DT	17-OCT-2003 (first entry)

DE DNA clone originating in barley containing SNP sequence #340.  
XX  
KW Barley; single nucleotide polymorphism; SNP; genotype-phenotype analysis

OS	Hordeum vulgare.
XX	
PN	MO2003057877-A1.

PD	17-JUL-2003.
XX	
PF	16-DEC-2002;

PR	20-DEC-2001; 2001JP-00387059.
PR	20-DEC-2001; 2001JP-00387131.
PR	20-DEC-2001; 2001JP-00403299.

PR 27-SEP-2002; 2002JP-0032751  
XX  
PA (UYN1-) UNIV JAPAN OKAYAMA.

PI Sato K, Takeda K, Kohara Y;  
XX  
DR WPI; 2003-587127/55.

PT Single nucleotide polymorphism sites in barley varieties and DNA  
PT sequences containing them for analysis and identification of barley  
PT varieties and production of barley transformants with desired

Pt		characteristics.
Xx	PS	Disclousure; SEQ ID XX; 284pp; Japanese.
Cc	CC	The present invention relates to oligonucleotide clones originating in barley ( <i>Hordeum vulgare</i> ) which contain single nucleotide polymorphisms (SNP). The oligonucleotides may be used for analysis of SNPs among barley varieties, identification of particular varieties and genotype-phenotype CC analysts, isolation of specific genes and creation of new varieties by transformation of barley varieties with them and production of new barley CC varieties with desired properties. The present sequence represents an CC oligonucleotide clone sequence featured in the specification. The CC sequence data for this patent did not form part of the printed WIPO at ftp.wipo.int/pub/published-pct-sequences
SQ		Sequence 289 AA;
	Query Match	45.5%; Score 615.5; DB 7; Length 289;
	Best Local Similarity	45.6%; Pred.No. 1,4e-63;
Matches	129; Conservative	47; Mismatches 66; Indels 41; Gaps 6
Oy	Db	10 SFIFGLNIVSFMWFLAPVPTFYKTKRKSSSEGYQAIPYMWALFSAGLLTYAVLRNA 69 :: ::::   :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: : 11 AFAEBLGNITSFTSLNPPIPFYIIFYKSKEGQSPPYYVALNSNMIMITYALKIGE 70 :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :
Oy	Db	70 YLIYSINGFGCAIELTVISLSPLFVAPRKSKIPTGMV-MLBEGALGWMPITYLAEGRH 128 :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: : 71 GLLTIRINAGCVIERVIILMYLVAPRRAKIIFTAKIVILLNVAGRGLTFLTLPFHGET 130 :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :
Oy	Db	129 RVMIVGTICCALINVAFPAFPUSIKMQVVKTSKVEMPTLSLFTLCATMFFYGPKKD 188 :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: : Dd 131 RVVSIGMICVGSPSCVPFAPAPLSIIEKRVIKTSVEVTMPSLTTLTISAVMWPLYGIID 190 :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :
Oy	Db	189 FYIAEPNLIGTFEVLQMQLVVNYDSKRIDEKSDPYREATSKSGYE-----I 238 :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: : Db 191 KYVALPNLIGLTFEGNIOMLVLYPMN-----ATPV-VASDAEGEKAMVPEADHV 241 :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :
Oy	Db	239 IINI-EDDNSDNA-----IQSMEXDPFR 260 :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: : Db 242 VINWGKADKSSCABVRPADVPRRCABAEEAAPGGQMVAVDFAR 284 :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :
R	RESULT 13	
ID	ABM73783	
XX	AC	ABM73783 standard; protein; 268 AA.  ABM73783;  17-OCT-2003 (first entry)
DE	Kw	DNA clone originating in barley containing SNP sequence #193.  Barley; single nucleotide polymorphism; SNP; genotype-phenotype analysis.  Hordeum vulgare.  WO200305787-A1.  17-JUL-2003.  16-DEC-2002; 2002MO-IB005403.  20-DEC-2001; 2001JP-00387059.  20-DEC-2001; 2001JP-00387131.  20-DEC-2001; 2001JP-00403299.  20-DEC-2001; 2001JP-00403300.  27-SEP-2002; 2002JP-00327515.  (UYNl-) UNIV JAPAN OKAYAMA.  Sato K., Takeda K., Kohara Y,

DR	WP1: 2003-587127/55.
XX	Single nucleotide polymorphism sites in barley varieties and DNA
XX	sequences containing them for analysis and identification of barley
PT	varieties and production of barley transformants with desired
PT	characteristics.
XX	
PS	Disclosure; SEQ ID XX; 284bp; Japanese.
XX	
CC	The present invention relates to oligonucleotide clones originating in
CC	barley ( <i>Hordeum vulgare</i> ) which contain single nucleotide polymorphisms
CC	(SNP). The oligonucleotides may be used for analysis of SNPs among barley
CC	varieties; identification of particular varieties and genotype-phenotype
CC	analyses; isolation of specific genes and creation of new varieties by
CC	transformation of barley varieties with them and production of new barley
CC	varieties with desired properties. The present sequence represents an
CC	oligonucleotide clone sequence featured in the specification. The
CC	sequence data for this patent did not form part of the printed
CC	specification, but was obtained in electronic format directly from WIPO
CC	at ftp.wipo.int/pub/published-pct-sequences
XX	
XX	Sequence 268 AA:
Query Match	45.5%; Score 615; DB 7; Length 268;
Best Local Similarity	49.6%; Pred. No. 1.5e-63;
Matches 121; Conservative	45; Mismatches 58; Indels 20; Gaps 4.
OY	10 SFIFGLGNISFVFWFLAPVPTFYKIKYKRSSEGYOAIPIYVWALFSAGLILYVAVLRKNA 69
Db	11 AFAFGGLGNISFSLAPITPTFRIFPKSSTBEFGSPYVVALFSAALMLFYALVKIGE 70
OY	70 YLIVSINGFCALIELTYISLEFLPAAPRKSKITFGWL-MLLELGAIGWMPITTYLIAEGSH 128
Db	71 GLITITINAAAGCVIEFTVYIIMLVVAPRKAKITFAKIVILLNVAGFGLIFLLTFAFHGET 130
OY	129 RVMIVGMICAIINAVFAAPISIKROYIKTSVSEFMPPTLSLFLTCATMMFFVGFFPKKD 188
Db	131 RVVSLGMCVGFSPVAFAPISIIIGVATIKTSVEVMPSPSLTTLTLSAVVWFLYGLIKD 190
OY	189 FYIAFPNIIGLFGIVQMLTFYVVKDSKRIDEXSDPYREATKSKEGVE-----I 238
Db	191 KYVALPNILGTFPGMIQWLYMFMN-----ATPV-VASDAKEGEAKMKVPADHIV 241
OY	239 IINT 242
Db	242 VINV 245
RESULT 14	
ID	AAAG42299
ID	AAAG42299 standard; protein; 294 AA.
XX	
AC	AAAG42299;
DT	18-OCT-2000 (first entry)
XX	
DE	Arabidopsis thaliana protein fragment SEQ ID NO: 52736.
XX	
KW	Protein identification; signal transduction pathway; metabolic pathway;
KW	hybridisation assay; genetic mapping; gene expression control; promoter;
XX	termination sequence.
XX	
OS	Arabidopsis thaliana.
XX	
PI	EPI033405-A2.
XX	
PD	06-SEP-2000.
XX	
PE	25-FEB-2000; 2000EP-00301439.
XX	
PR	25-FEB-1999; 99US-0121825P.
PR	05-MAR-1999; 99US-0123180P.
PR	09-MAR-1999; 99US-0123548P.

PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130649P.  
PR 23-APR-1999; 99US-0130610P.  
PR 23-APR-1999; 99US-0130891P.  
PR 28-APR-1999; 99US-0131499P.  
PR 30-APR-1999; 99US-0132048P.  
PR 30-APR-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 06-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142277P.  
PR 13-JUL-1999; 99US-0143342P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144328P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 28-JUL-1999; 99US-0145986P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154799P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.





**This Page Blank (uspio)**